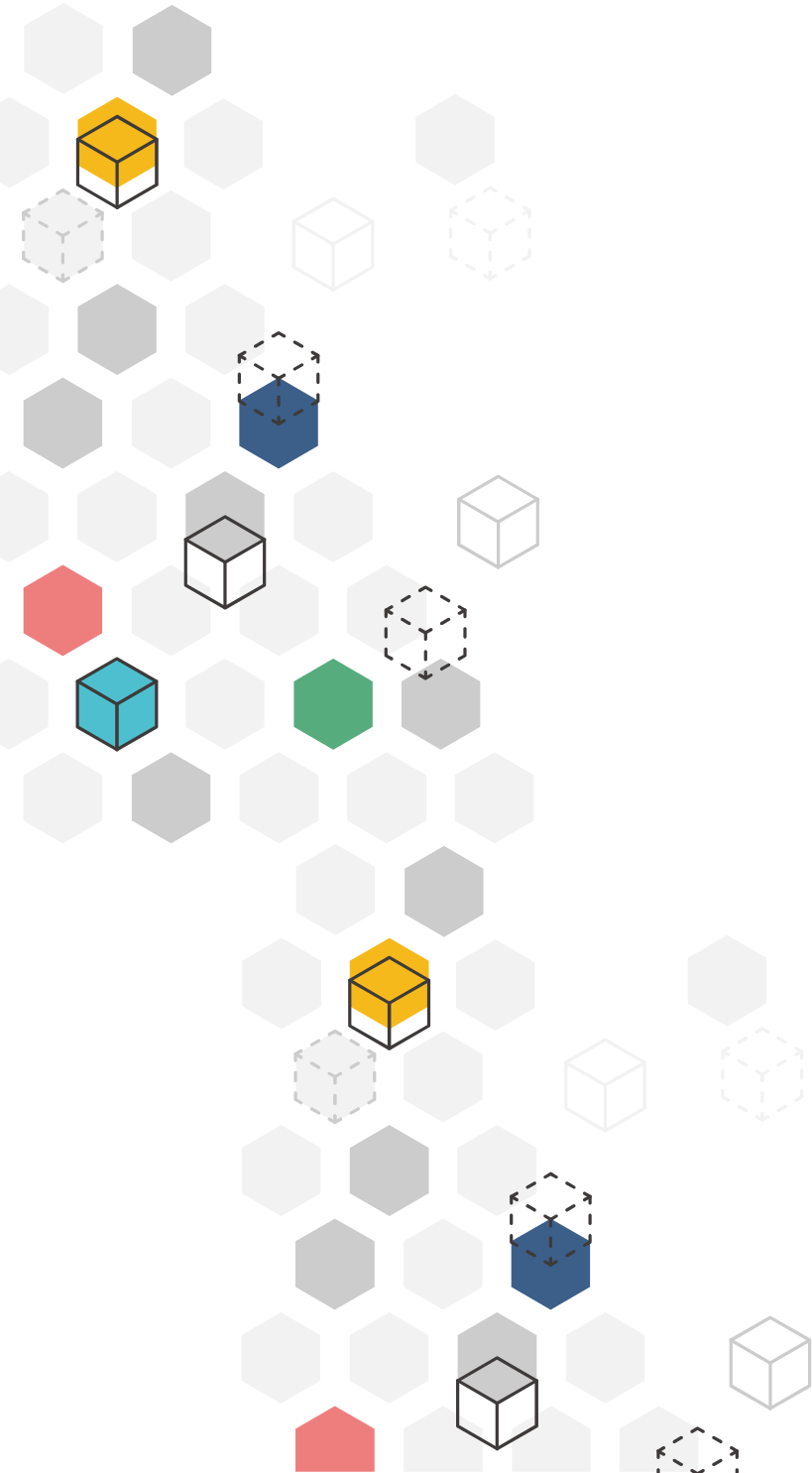
2019
정보보호 데이터바우처 성과사례집

Kdata 한국데이터산업진흥원
KISIA 한국정보보호산업협회



2019
정보보호데이터바우처 성과사례집

CONTENS

한국정보보호산업협회
I. 기관소개

한국정보보호산업협회(KISIA)

정보보호산업진흥법 제 24조에 의해 설립된 법정법인입니다. KISIA는 1998년 설립 이래 한국정보보호산업의 건전한 발전과 국가산업 전반의 정보보호 수준 제고를 위한 사업 환경 조성 및 상호협력 도모를 위해 끊임없이 노력해왔습니다. KISIA의 주요 사업은 다음과 같습니다.

한국정보보호 산업협회(KISIA) 주요 사업

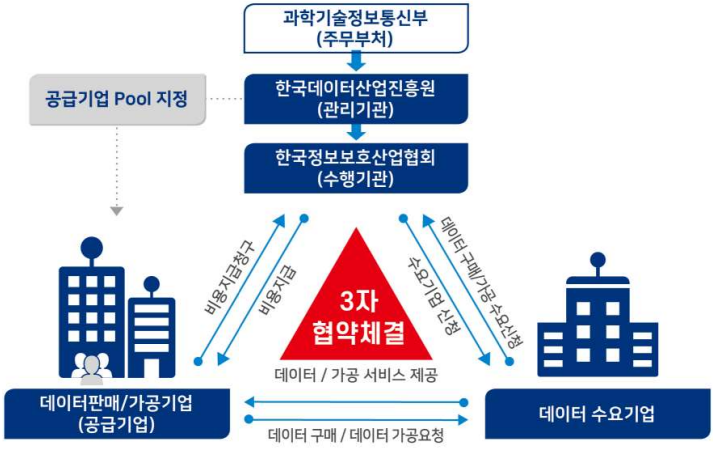
- 정보보호 관련법 제도의 효율적인 개선방안 건의
- 산업체 애로 사항 및 회원사 의견 수렴 및 해결
- 산업 현황 파악 및 정책 수립을 위한 시장동향 및 통계조사
- 유관기관과의 일원화된 업무로 인한 신속한 정보 공유
- 국내·외 전문전시회 참가 지원 사업 및 세미나 지원
- 기술 개발을 위한 산·학·연 연구 및 전문 인력 양성
- 정부 및 관계 기관과의 위탁 조사연구 수행

데이터바우처 지원사업
Ⅱ. 소개

데이터바우처 지원사업

데이터바우처 지원사업은 데이터 활용에 어려움을 겪는 정보보호 · 일반 중소, 벤처기업, 소상공인, 1인 창조 기업에게 데이터 구매 · 일반가공 · AI 가공에 필요한 금액을 바우처로 지원하는 사업입니다. 본 사업은 과학기술정보통신부가 주최하고, 한국데이터산업진흥원이 주관하며 한국정보보호산업협회는 수행기관으로 참여하고 있습니다. 공급기업의 등록 및 관리는 한국데이터산업진흥원에서 담당하고 있으며, 한국정보보호산업협회가는 수요기업의 접수 및 관리업무를 수행하고 있습니다.

[그림] 데이터바우처 지원사업 구조도



1-1.
AI가공

나루씨큐리티

Making Sense of the Unknown
NARUSECURITY

기 업 명	나루씨큐리티
참 여 부 분	AI 가공
데이터 활용	네트워크 플로우 데이터 가공
공 급 기 업	엘렉시
성 과	내부망 정보보호 위협 AI 탐지 모듈 개발

AI를 활용하여 내부망 네트워크의 위협을 탐지하다

최근 사이버공간의 화두는 방화벽, 안티바이러스 등과 같이 내부망에 침투하여 지속적인 활동을 통해 내부 민감정보를 탈취하거나 시스템을 파괴하지만 방어체계에는 탐지되지 않는 고도의 사이버 공격행위입니다. 나루씨큐리티는 이를 탐지하기 위한 혁신적 사이버 공격 가시화 및 데이터기반의 의사결정지원 체계를 제공합니다. 다수 체계의 연동을 통한 연계 분석 방식과 달리 하나의 시스템을 통한 데이터 수집과 분석으로 내부망에서 은밀하게 움직이는 공격자의 움직임을 파악하고, 인과관계를 분석해서 예방체계를 우회한 공격을 명확하게 식별합니다.

나루씨큐리티는 미국 국제특허를 획득한 고도의 데이터 분석 기술을 기반으로 복원력 있는(Cyber Resilience) 네트워크를 구성하도록 하며, 이를 통해 사이버공격으로부터 발생하는 비즈니스 연속성의 침해를 원천적으로 방지하는 솔루션을 제공합니다.

주요 제품으로는 커넥텀(내부망 보안), 사이버배틀필드(사이버 침해사고대응 훈련 시스템)가 있습니다. 데이터바우처 지원사업을 통해 자사의 내부망 정보보호 위협탐지 솔루션 '커넥텀'을 AI화 하는데 소요되는 시간 및 경비를 절감할 수 있었습니다.

[사진] 나루씨큐리티 사무실

사업에 지원하게 된 동기

빅데이터를 활용하여 내부망 내 악성행위를 탐지할 수 있을까?

내부망 은닉 공격 솔루션을 보유한 나루씨큐리티는 증가하는 내부망 위협에 따라, 진보된 탐지 기술이 필요하다고 생각했고, 최근 급증하는 APT 공격 및 내부망 침투 공격에 각 기업이 대응하기에는 기존 정보보호 솔루션으로는 역부족이라고 생각했습니다. 이러한 상황을 개선하고자 각 기업 정보보호 담당자에게 내부망 위협에 대한 영감을 줄 수 있는 솔루션과 진보된 탐지 장비가 필요하다고 느꼈고, 기존의 솔루션이 가지고 있는 한계점을 극복하면서 정보보호 담당자가 한눈에 볼 수 있는 기업 내부망의 위협 및 현황을 시각화한 데이터가 있다면 관리가 수월하지 않을까 생각했습니다. 그래서 데이터바우처 지원사업에 지원하게 되었고, 진보된 솔루션 개발에 필요한 자원 및 서비스를 제공받을 수 있었습니다.

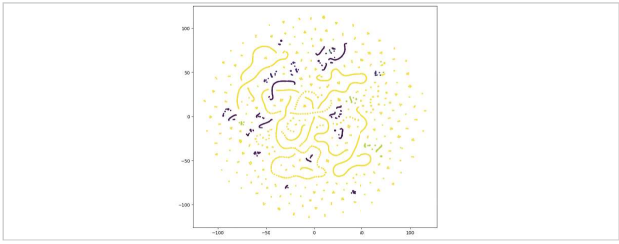
사업 진행과정

무의미한 텍스트에서 유의미한 자료가 되기까지

기존 솔루션에서 생성된 네트워크 플로우 정보를, 데이터 가공이 용이하도록 정규화 과정을 거친 후, 시계열 클러스터링 방법론을 적용하여 검증하는 작업을 실시하였습니다.

[그림] 가공데이터의 형태

예측값	예측된 Label																
실제값	AMA-log	Google	AMA-au	AMA-blu	AMA-lam	AMA-bl2	Avast	Acronis	Symantec	Carpathia	Oracle	Github	Fortinet	Alibaba	Uunet	Mcafee	Sejong
AMAZON log	9027	5	7	2	4	1	674	82	0	491	0	1418	118	188	42	41	0
Google android	611	9912	4	5	5	0	29	28	6	35	11	1309	49	72	7	8	9
AMAZON auth	0	3	12024	4	35	0	0	1	0	0	0	2	8	10	9	5	0
AMAZON bluecoat	1	6	9	12034	2	1	17	7	1	0	0	7	11	1	3	0	0
AMAZON lam	2	16	527	0	11406	0	1	6	0	0	0	3	121	4	7	6	0
AMAZON bluecoat2	2	0	0	4	0	12092	1	1	0	0	0	0	1	0	0	0	0
AVAST slc	873	1	0	4	0	1	4667	4950	7	182	0	788	469	151	0	3	0
ACRONIS cloud	1	0	0	1	0	0	0	11725	0	236	0	3	111	6	1	17	0
SYMANTEC bluecoat	0	0	0	5615	1	12	0	0	0	0	0	6457	10	1	0	1	0
CARPATHIA	136	0	0	2	0	0	11	64	0	9770	0	151	932	695	339	1	0
ORACLE dyndis	0	1	4	1	0	0	0	0	0	2	12085	3	1	3	0	0	0
GITHUB numerabjs	1508	9	4	6	1	5	223	32	2	196	6	9875	127	43	15	48	0
FORTINET	15	17	339	7	69	0	83	2628	8	3617	3	27	5224	21	40	2	0
ALIBABA-US	81	70	24	1	13	2	331	48	6	103	0	123	67	11085	132	14	0
UUNET	20	2	4	3	5	3	6	1523	0	6409	0	42	1354	101	2627	2	0
MCAFEE sercie	66	2545	688	12	1666	0	11	239	0	12	1	114	81	287	5	6374	0
SEJONGTELECOM time	0	12098	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4



데이터 활용 방법은 아래 그림과 같습니다. 내부망 정보보호 위협탐지 장치인 커넥텀이 빅데이터 분석을 통해 내부망에 탐지되지 않은 정보보호 위협을 탐지 및 분석합니다.

※ 아래 그램에서 커넥텀의 분석영역은 노란색으로 표시

[그림] 데이터 활용 방법 개념도

위 그림에서 커넥텀에 의해 탐지 및 분석된 악성행위는 명령제어채널, 내부망이동, 정보유출, 내부장악 4가지 징후로 이 중 가장 첫 단계인 명령제어채널 통신 분석의 자동화를 통해 기존의 정보보호 체계를 우회하여 내부망에 은닉하고 있는 악성코드의 활동을 탐지하고 대응이 가능 하도록 하였습니다.

일반적인 크기의 네트워크(대략 1000대 이상의 호스트로 구성된)에서는 대략 2000여개 이상의 지속통신이 발생하며, 이중 80%인 1600여개는 망사업자 번호 및 기존 인텔리전스에 의한 자동분류가 가능하나 나머지 400여개의 통신은 사업장 별 인적 분석수요가 발생하게 됩니다. 따라서 나머지 400개의 통신에는 본 사업을 통해 개발한 통신자동분류 기능을 적용했습니다.

사업의 성과

데이터바우처 지원사업으로 AI 탐지 모듈 개발을 앞당기다!

본 사업에서 획득한 통신자동분류기능을 활용하여 90% 이상의 통신데이터 자동분류를 가능하게 하여 업무생산성을 향상시키고 수동 분석에 따르는 비용을 절감했습니다. 향후에는 IoT, ICS 등 특정 목적기반 네트워크에서 발생하는 통신의 학습을 통해 시그니처에 의존하지 않는 인공지능기반 악성코드 탐지 백신 개발이 가능하도록 할 예정입니다.

현재 몇몇 대형사업장을 제외하면 충분한 정보보호 인력을 갖추기 어려운 것이 사실입니다. 그래서 나루씨큐리티는 이번 사업을 통해 획득된 데이터를 활용하여 시큐리티 데이터 분석 사업에 활용하여 보다 낮은 가격으로 양질의 보안 분석 서비스를 제공받을 수 있도록 하려고 합니다. 또한 향후, AI 분석 기능이 탑재된 은닉 공격 탐지 솔루션을 출시하고, 데이터바우처 사업을 통해 개발된 AI 분석 기능을 자사 솔루션인 커넥텀에 탑재하여 더 강력한 자사 솔루션을 시장에 선보일 예정입니다. 또한 AI 분석 기능을 활용한 내부망 원격 관제 서비스도 출시할 계획입니다.

사업에 대한 의견

내부망 네트워크에 대한 위협은 꾸준히 증가하고 있는 상황이지만, 중소기업의 정보보호 솔루션에 대한 투자는 미비한 실정입니다. 데이터바우처 지원사업을 통해 전문기업에 소속된 AI 전문가의 데이터가공 서비스를 받을 수 있었고, 비용도 절감할 수 있었습니다. AI 솔루션을 개발하려고 하는 저희와 같은 정보보호 중소기업들에게 꼭 필요한 제도라고 생각합니다.

또한, 개별 기업이 보유한 데이터는 충분히 가치 있고, 무한한 잠재력을 가지고 있다고 생각합니다. 무한한 잠재력을 가지고 있는 데이터가 가공되어 활용도가 높아진다면, 잠재력에서 머물지 않는 실질적인 가치를 창출할 수 있을 것입니다. 가공되지 않은 데이터는 원석일 뿐이지만, 데이터바우처 지원사업을 통해 가공하고 다듬어진다면 빛나는 보석이 될 수 있습니다. 하지만 데이터바우처 사업에 참여하여 사업성과를 기대만큼 달성하기 위해서는 내가 왜 이 사업에 참여하는지에 대한 '뚜렷한 목적의식'이 필요합니다.

내가 무언가를 만들고 싶지만 어떻게 만들지 그 과정이 어렵거나, AI 분석과 같은 빅데이터 분석을 통해 데이터를 더 단단하게 만들고 싶을 때 유용한 기회가 될 것입니다.

1-1.
AI가공

에프원시큐리티

F1Security

기 업 명	에프원시큐리티
참 여 부 분	AI 가공
데이터 활용	인공지능 악성코드 식별, 분류 등
공 급 기 업	호서텔넷
성 과	AI 가공 서비스를 통한 기존의 악성코드 탐지 성능 개선

웹 악성코드 인공지능 분석 프레임워크로 악성코드 분석의 효율성, 웹 유포 탐지의 정확성을 높이다.

에프원시큐리티는 과학기술정보통신부 지정 기반 시설 정보보호 전문 서비스 기업으로 정보보호컨설팅과 웹보안 솔루션, 연구개발, 악성코드 분석 대응 및 웹 취약점 점검 서비스 등을 제공하고 있습니다. 인공지능이 필요한 최적의 타이밍에 데이터바우처 지원을 받아 인공지능을 접목한 악성코드 분석탐지 연구 개발에 박차를 가할 수 있게 되었습니다. 인공지능을 접목하여 악성코드 분석 기능을 개발하고, 탐지 성능을 개선 할 수 있었습니다.

[사진] 에프원시큐리티 임직원

사업에 지원하게 된 동기

인공지능을 이용하여 악성코드 분석 자동화 및 탐지 정확도를 개선하고 악성코드 분석과 탐지를 위한 인공지능 피처 가공의 필요성을 느끼다

인공지능 기술은 보안시장에 새로운 기회를 제시, 빅데이터를 활용하여 사람이 찾아내지 못하는 특성을 찾아 새로운 악성코드 탐지 기술 탄생의 기회를 제공합니다.

인공지능을 통한 중소기업의 기술적 도약 및 악성코드 웹 유포 방식이 변화함에 따라 새로운 웹 유포 대응이 필요했습니다.

기존에는 악성코드를 유포하는 웹 주소 즉 유포지 URL을 이용하여 유포를 탐지해 왔으나, 최근에는 유포지 URL의 재사용율이 떨어지고 시간차 공격, 이메일 APT 등 유포방식이 지능화 되고 있습니다.

이에 웹 사이트에서 다운로드 되는 파일의 악성여부 점검을 통한 유포 탐지 방안 필요하다고 판단하였습니다. 인공지능을 이용한 악성코드 탐지를 위해서는 데이터바우처 지원사업을 통해 다양한 피처를 식별하고, 인공지능 데이터 분석을 통하여 최적화된 피처를 구성해야 한다고 생각했습니다.

그동안 인공지능 탐지를 위하여 악성코드를 수집하고 인공지능 연구 인력을 채용하는 등의 투자를 하였으나, 중소기업이 빅데이터 인공지능 개발의 다양한 연구 인력을 확보하기는 쉽지 않았습니다. 데이터바우처 지원을 통해 데이터 분석, 인공지능 가공서비스를 지원받아 악성코드 탐지 피처 가공을 통한 악성코드 탐지 성능을 고도화 할 수 있었습니다.

사업 진행과정

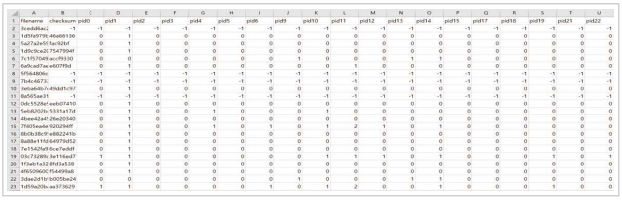
인공지능 악성코드 피처를 가공 및 분류하여 악성코드를 분석하다.

데이터의 종류 (바이너리 파일)		
실행파일의 binary labeling 정상파일, 악성파일	실행파일의 multi labeling 정상파일, 랜섬웨어, 스파이웨어, 다운로더 등	정상 10만, 악성 10만

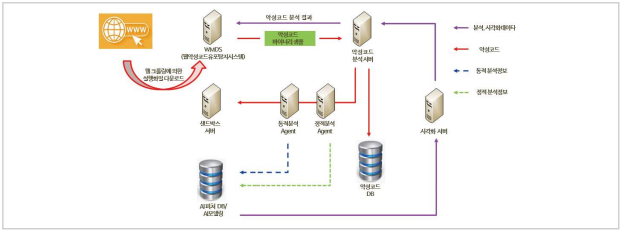
데이터 가공 방식	
① 추출	• 샌드박스로부터 동적 악성행위 정보 추출 • 악성코드 PE로부터 정적 피처 추출
② 전처리	• 악성코드로부터 추출한 정보의 피처의 분리, 조합, 코드화 • API/Opcode sequence, ngram 처리 • API/Opcode 의 코드색상화, 동일크기 image 등
③ 데이터 분석	• 피처 선별 및 가중치, 악성코드 유형 등을 수동분석 및 인공지능 결과를 토대로 검증 분석 등
④ 인공지능 악성코드 식별, 분류	• 악성코드 식별 및 유형 분류, 라벨링
⑤ 시각화	• 연관관계 시각화

추출 및 가공된 데이터를 이용하여 인공지능 학습모델 개발, 학습모델을 이용하여 정상과 악성을 판별합니다. 실행 파일의 추출 및 가공 데이터, 연관분석 및 시각화를 활용해서 악성코드를 분석합니다.

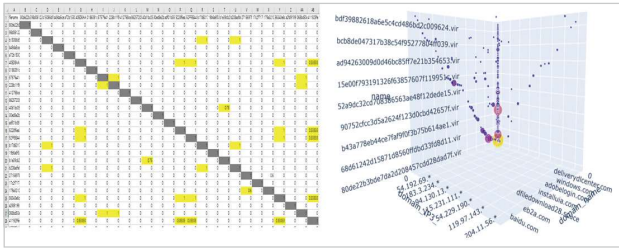
[그림] 컴파일 정보를 담고 있는 리치헤더 피처



[그림] 데이터 활용 방법에 대한 개념도



[그림] 문자열 비교를 통한 유사도 분석과 3차원 그래프로 시각화된 모습



사업의 성과

인공지능 악성코드 분석 및 탐지 성능을 개선하다

기존의 악성코드 탐지 성능을 개선하고 연관분석 및 시각화를 구현할 수 있게 되었습니다. 또한, 악성코드에 대한 분석과 탐지 방안에 대한 추가적인 아이디어를 확보할 수 있었습니다. 추후에는 인공지능 악성코드 분석 및 탐지 특허를 상용화할 예정이며, 산학협력 및 연구 과제를 통한 인공지능 보안 기술의 고도화를 위해 지속적으로 노력할 계획입니다.

사업에 대한 의견

인공지능 기술과 빅데이터 가공 개발, 빅데이터 분석 등의 전문 기술과 인력을 지원받아 단기간에 원하는 성과를 거둘 수 있었습니다. AI 신기술에 대한 도약의 기회가 되었고, 솔루션의 기능 개선에 큰 도움이 되었습니다.

인공지능은 4차 산업혁명의 핵심 기술입니다. 데이터바우처 지원사업을 통하여 기업 내부의 인공지능 기술의 내재화와 기업의 빅데이터 활용 방안을 모색해 보시면 좋을 것 같습니다.

1-1.
AI가공

와치포인트

WatchPoint

기 업 명	와치포인트
참 여 부 분	AI 가공
데이터 활용	내부자 행위 데이터셋 AI 학습셋 가공
공 급 기 업	띵스파이어
성 과	AI 내부정보유출관제 솔루션 개발

내부정보 유출! 그게 가능해?
WatchPoint-UEBA와 함께 하면 불가능합니다.

와치포인트는 사내 내부정보 유출을 관제하기 위한 솔루션 개발을 주요 사업으로 진행하고 있는 보안 전문기업이며, 주요 제품으로는 WatchPoint-UEBA가 있습니다. 그 외 실시간 전력량 분석을 통한 UEBA(User Entity Behavior Analytics) 개념의 EMS (Energy Management System) 솔루션 사업과 블록체인 관련 서비스 개발 등을 진행하고 있습니다.

내부정보 유출관제 솔루션은 기업의 민감한 정보를 취급하기 때문에 테스트 데이터 확보에 어려움을 겪고 있었으나, 데이터바우처 지원사업으로 제공받은 데이터로 신규 솔루션 개발에 많은 도움이 되었습니다.

[사진] 와치포인트 사무실

사업에 지원하게 된 동기

내부정보 유출관제 솔루션에 AI를 적용하고 싶은데 관련 데이터와 기술을 확보할 수 없을까?

AI기술이 화두가 되면서 많은 기업이 주요 시스템 및 솔루션에 AI기술을 도입하고 있지만, 보안 분야는 정작 기술 개발을 위한 데이터 확보조차 쉽지 않은 것이 현실입니다 "대기업에서 내부정보유출 통합관제 시스템을 구축하는데 무조건 AI 기술을 적용하라고 요구하였죠. 상당히 민감한 데이터를 다루는 솔루션이다 보니 관련 기술개발을 위한 데이터 확보도 불가능하고 그러다 보니 프로젝트 수주도 실패했습니다.

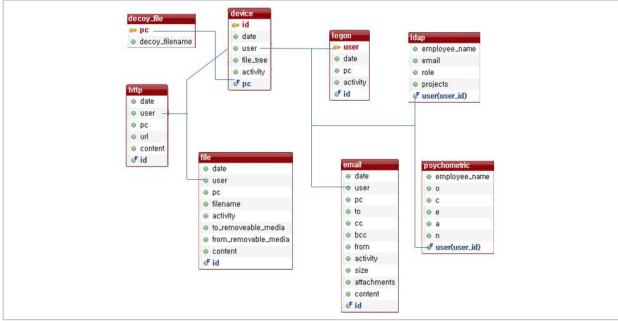
그리고 기술이전 행사도 많이 참석했는데 방법을 못 찾았습니다. 그러던 중 우연히 데이터바우처 지원사업에 참여할 수 있는 기회가 되어 관련 데이터를 확보할 수 있었습니다. 이후 우선 머신러닝 기능을 먼저 도입했고 관련 프로젝트도 수주했습니다. 바우처 지원 사업은 우리 회사에 정말 필요했던 데이터와 AI기술 확보를 가능하게 해주었습니다. 내년에는 제품을 더욱 고도화시켜 사업을 확장시키려고 계획 중입니다.

사업 진행과정

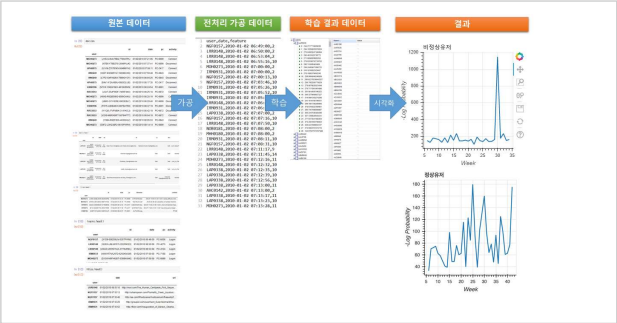
내부정보유출관제 솔루션 WatchPoint-UEBA에 AI 갑옷을 입히다

카네기멜론대학교의 내부자 행위 데이터셋을 AI 학습셋으로 가공하는 과정을 거쳐, AI 내부 유출 관제 시스템에 탑재하여 관련 기술을 개발하였습니다.

[그림] 원천데이터 구조도

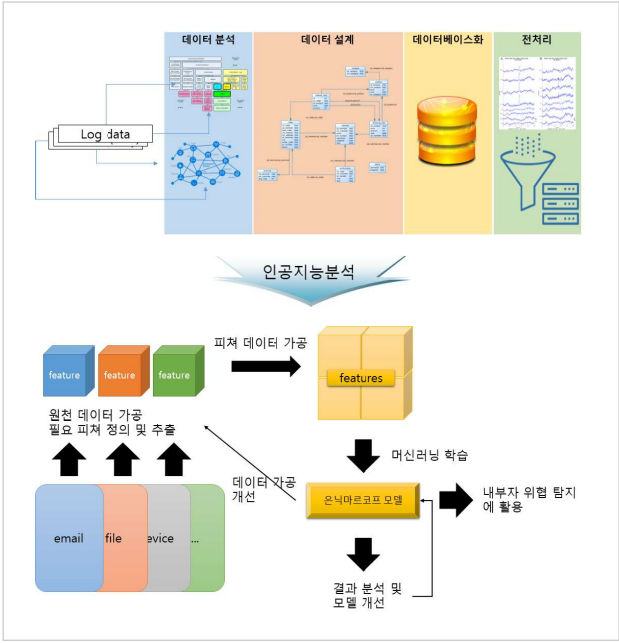


[그림] 원천데이터 가공 구조



기존에는 내부정보 유출관제 솔루션을 개발하기 위한 개발 데이터를 확보하는 것이 불가능했습니다. 관련 프로젝트를 수행하더라도 워낙 중요한 데이터라서 절대 외부 반출이 불가능했습니다. 그러다 보니 개발을 하면서도 제대로 했는지 확인도 못해보고, 눈감고 개발하는 것과 비슷한 상황이 되었습니다. 하지만 데이터바우처 사업에 참여하면서, 개발 데이터를 확보해 실제 데이터를 보면서 작업을 하니 이제는 제대로 된 솔루션을 개발할 수 있다는 자신감이 듭니다.

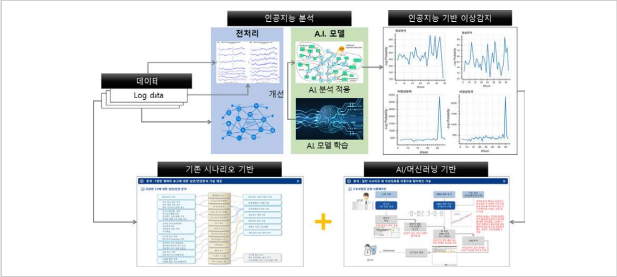
[그림] 데이터 활용 방법 개념도



사업의 성과

데이터바우처로 내부정보유출관제 신제품을 출시하다!

[그림] 데이터바우처 지원 이후 사업 형태



데이터바우처 지원사업을 통해 AI 기술이 적용된 신제품을 개발하였습니다. 기존에는 시나리오 기반의 제품이다 보니 세세하게 시나리오를 설정하기도 너무 힘들고, 특히 사전 예측 자체가 불가능했습니다. 이제 양손에 칼을 하나씩 든 기분입니다. 시나리오 기반과 AI 기술을 적절히 활용하면 더욱 효율적인 관제가 가능할 것이라 생각됩니다. 보안은 절대적인 것이 없습니다. AI가 좋은 경우도 있지만 시나리오 기반이 효율적인 경우도 많습니다. 저희는 두 가지 기술 모두를 확보한 셈이죠.

내년부터 본격적으로 AI기술을 고도화해 나갈 계획입니다. 하지만 하루아침에 AI가 완벽해질 수는 없을 것입니다. 특히 사람 행동을 관리하고 예측하는 제품으로, 은밀하게 행해지는 보안 영역이니 어련하겠습니까? 데이터바우처 지원사업 덕분에 올해 작은 프로젝트도 수주했습니다. 이를 기반으로 추후에는 제대로 된 국산 솔루션을 만들고 싶습니다. 기존 제품의 장점과 AI 기술의 강점을 접목하면 제대로 된 제품이 되지 않을까요?

사업에 대한 의견

데이터는 참 많은 것을 보여줄 수 있다고 생각합니다. 다만 제대로 된 데이터를 확보하기가 너무 어렵습니다. 보안 데이터는 더욱 어렵고 중소기업에게는 더 어려운 일 입니다. 공공데이터 제공하는 곳을 살펴보았는데, 데이터는 있지만 제대로 활용할 수 있는 것이 많지 않았습니다. 이왕 제공할 것이라면 쓸 만한 데이터를 제공해야 한다고 생각합니다. 데이터바우처 지원사업에서 조금 더 가치 있는 데이터를 제공할 수 있도록 보완해주시면 데이터바우처 지원사업이 더욱 활성화되지 않을까 생각됩니다. 앞으로의 산업 생태계는 한 기업이 모든 것을 담당하고, 개발하는 시대에서 서로 협력하고, 이종 간의 산업 간에도 콜라보레이션이 가능해야 하는 시대로 넘어가고 있습니다. 따라서 현재의 서비스 시스템을 보다 지능화 및 고도화 하고자 하는 기업들은 데이터바우처 지원사업에 적극적으로 참가하여 다양한 기업들과 협력할 수 있는 기회를 만들기를 추천 드립니다.

1-2.
일반가공

세인트시큐리티

SAINT SECURITY
malwares.com™

기 업 명	세인트시큐리티
참 여 부 분	일반가공
데이터 활용	악성메일 시그니처 정보 가공
공 급 기 업	지란지교시큐리티
성 과	악성코드에 대한 정보 제공 확대 및 문서 형태 악성코드의 유해컨텐츠 무력화 기술 개발

APT 공격 사전 탐지 및 방어를 위한 스팸 데이터 활용 및 유해 컨텐츠 무력화 기능 구현

세인트시큐리티는 창업한 지 15년 되는 악성코드 수집 및 분석 관련 전문 업체로써 각종 악성코드 관련 인텔리전스 제공 서비스를 malwares.com을 통해서 제공하고 있습니다. 최근 들어 머신러닝 기반 엔드포인트 보안 제품인 MAX와 네트워크 기반 제품인 MNX를 개발하여 시장에 선보이고 있습니다. 해외 경쟁 서비스 대비 차별성 있는 정보를 제공하고 싶었는데, 지란지교시큐리티의 스팸 관련 정보를 통해서 악성코드와 스팸 관련 정보를 함께 프로파일링 하여 제공 할 수 있어, 정보의 값어치를 높일 수 있었습니다.

[사진] 세인트시큐리티 대표님

사업에 지원하게 된 동기

해외 기업에서 제공하는 서비스와는 차별화 된 경쟁력 있는 정보를 국내에서 생성할 수 없을까?

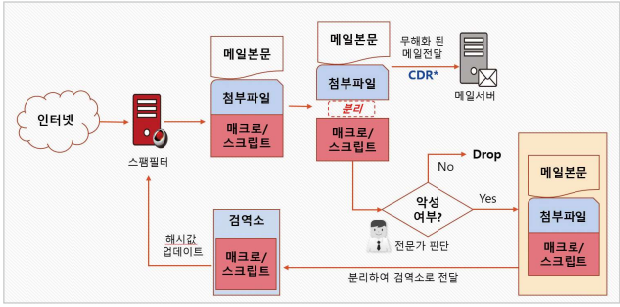
새로운 데이터를 합쳐서 새로운 서비스 혹은 제품을 개발 했을 때 매출이 발생한다는 보장도 없고, 그렇다고 무턱 대고 투자할 수도 없는 상황이었습니다. 또한, 자사의 서비스만 투자를 해서 될 일이 아니라 타사의 서비스와 정보도 필요했던 상황이었는데, 이런 상황에서 몇 천 만원의 투자를 하는 것이 쉽지 않았습니다. 그래서 국가에서 지원 해주는 데이터바우처 지원사업을 통해 지란지교시큐리티와 함께 사업을 수행할 수 있게 되었습니다.

사업 진행과정

기존 데이터를 취합 및 가공하여 해외 서비스와 차별화되는 국내 전용 데이터 서비스를 새롭게 만들자!

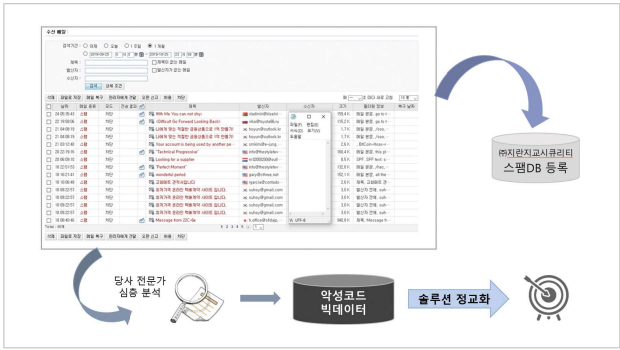
네트워크 트래픽을 통해 수집되는 각종 메일 관련 정보를 수집하고 수집한 메일 중에 첨부 파일이 있는 경우 해당 첨부 파일의 위험도를 분석했습니다. 첨부 파일 중에는 실행 파일이 있을 수 있고, 대부분은 문서 형태가 많은데, 최근에는 문서 형태의 파일로 제로데이 취약점을 이용한 공격이 많이 발생하고 있습니다. 이러한 경우 해당 문서를 수집하고 위험도를 식별한 다음 해당 문서에 포함된 각종 위험 요소들을 제거해야 합니다. 그리고 수집된 악성코드 및 메일 정보를 분석하여 누가 누구에게 어떤 제목과 내용으로 송신했는지 프로파일링 하여 해당 정보를 가공했습니다. 가공된 정보는 malwares.com이나 지란지교시큐리티에 전달되어 추후 새로운 위협 정보를 식별하는데 다시 활용될 예정입니다.

[그림] 데이터 활용 방법에 대한 개념도 ①



* CDR : Content Disarm & Reconstruction (콘텐츠 무해화)

[그림] 데이터 활용 방법에 대한 개념도 ②



사업의 성과

국내에 최적화 된 정보 생성과 제공으로 해외 보안 업체의 국내 시장 장악을 방지하고, 이를 기반으로 오히려 차별화되는 서비스와 정보로 해외 시장에 진출 할 수 있는 교두보를 마련하다.

기존에는 악성코드에 대한 위협 정보만을 제공했지만 본 사업을 통해 가공된 데이터를 통해서 해당 악성코드의 유포 경로 정보를 함께 제공 할 수 있게 되었습니다.
또한, malwares.com을 통해 수집된 다양한 문서 형태의 악성코드에 포함된 유해 컨텐츠를 식별하고 이를 무력화 할 수 있는 기술을 개발하였습니다.
향후에는 차별화 된 정보를 기반으로 3~6개월 정도 시범 서비스로 기존 고객들에게 먼저 데이터를 제공해서 반응을 살펴본 다음, 일반 사용자에게 공개 하는 방식으로 진행할 예정입니다.
만약에 필요하다면 추가 가공을 통해 좀 더 데이터를 다듬어야 할 필요성이 있다고 생각합니다. 또한, 스팸 메일 수·발신에 대한 정보를 포함하고 있다 보니 개인 정보 보호에 대한 이슈도 있어서 스팸 발신자에 대한 정보도 개인 정보에 포함되는 지 확인하여 서비스를 할 필요성도 있습니다.

사업에 대한 의견

국내에서 데이터의 값어치에 대해서 가이드를 제시하고 이에 대한 필요성을 충분히 언급할 수 있는 사업이라서 좋았습니다. 하지만 공급기업의 독자적인 사업이 되지 않도록 노력 할 필요성이 있고, 수요기업이 새롭게 만든 데이터로, 수요기업이 공급기업이 되는 선순환을 통해 계속해서 발전 하고 데이터의 품질을 향상시킬 수 있는 좋은 구조가 만들어 질 수 있었으면 좋겠습니다.

IT 관련 사업을 하는데 있어 가장 필요한 것이 자금과 사람입니다. 결국 그 자금과 사람을 통해서 사업에 필요한 여러 정보를 가공하고, 가공된 정보를 활용하게 되는데 본 사업을 통해서 그 가공된 정보를 사업에 활용 할 수 있어서 좋았습니다. 데이터가 필요하고 관련된 사업을 하시는 분은 6개월에서 길게는 1년의 시간을 단축 시켜 Time to market을 실현할 수 있는 기회이니 주저 하지 말고 도전하여 사업을 수행 하는데 도움이 될 수 있기를 기원합니다.

1-2. 일반가공

엔에스에이치씨

기 업 명	엔에스에이치씨
참 여 부 분	일반가공
데이터 활용	악성코드 분석 데이터 대시보드 가시화
공 급 기 업	망고클라우드
성 과	모바일 악성코드 데이터 위험 가시성 확보

데이터 가공으로 머신러닝 악성코드 분석 포털의 디지털 가시화 기능을 완성하다.

엔에스에이치씨는 국·내외 최신 보안 위협요소를 파악하여 Anti-Virus, 악성코드 분석, 화이트 해킹, 모의 취약점 진단 서비스 등 다양한 보안 솔루션을 제공하는 R&D 기반의 정보보안 전문 기업입니다. 데이터바우처 지원사업으로 당사의 R&D 사업 아이템인 '머신러닝 악성코드 분석 통합 서비스'의 테스트베드 환경 및 상용화의 토대를 마련할 수 있었습니다.

사업에 지원하게 된 동기

악성코드 데이터 분석 결과를 이해하기 쉽고, 간결하고, 사용자가 편리하게 파악 할 수 있는 대시보드가 있으면 어떨까?

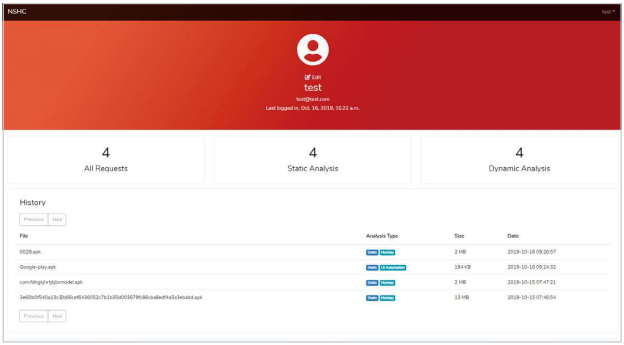
자사가 개발 중인 악성코드 탐지 및 분석 서비스는 분석 결과의 Data Visualization 구현이 포탈 기능 개발로서 수행되어야 하는데, 자사에는 마땅한 수행 가능 인력이 없는 상태였기 때문에 지원 신청을 하게 되었습니다. 자사가 개발하는 머신러닝 기반 악성코드 분석 플랫폼은, 사용자의 편의성을 기반으로 한 분석 결과 및 통계 정보를 악성코드 연구자들에게 유용하게 제공합니다. 그리고 이를 활용하는 보안 응용 서비스를 발굴하는 것이 향후 저희 솔루션의 상용화 요건이라고 생각합니다. 그러나 그 전 단계로 테스트베드 포탈 환경을 구축해야하는데 아쉽게도 내부에는 이러한 데이터 가시화 기능 개발 경험자가 없었습니다. 하지만 데이터스토어 홈페이지를 통해 적합한 가공기업을 찾을 수 있었습니다.

사업 진행과정

악성코드 데이터의 가시화로 분석결과를 한눈에!

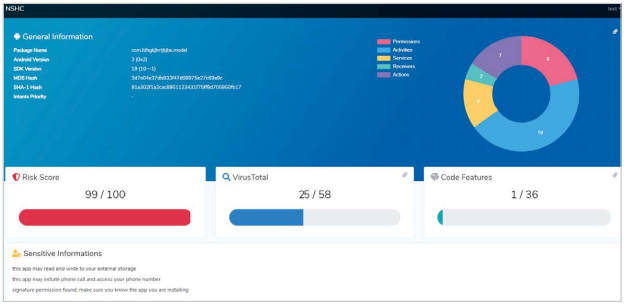
자사에서 지속적으로 크롤링한 20MB 이하의 우회기술이 적용된 다양한 악성코드 샘플을 1주당 천개씩 확보하여 가공기업에게 제공하였습니다. 가공기업에서는 이 샘플을 전처리하여 개발 서버 내부에 저장한 다음, 악성코드의 데이터 및 특성정보를 분석 시스템을 통해 정적 분석하고, 그 결과를 NoSQL 형태로 데이터 저장소에 저장했습니다. 그 다음에는 분석 시스템에서 수행되는, 실제 안드로이드 환경의 동적 분석 과정으로 악성코드 파일을 실행하여 안드로이드 환경의 변화를 측정합니다. 이 과정에서 나오는 feature값, 위협 행위 정보 유형, 탐지 정보 등을 종합적으로 분석하여 그 결과를 대시보드 형태로 해서 데이터 가시화 기능을 구현하였습니다.

[그림] 사용자 분석 이력 및 과거 악성코드 관련 정보 조회



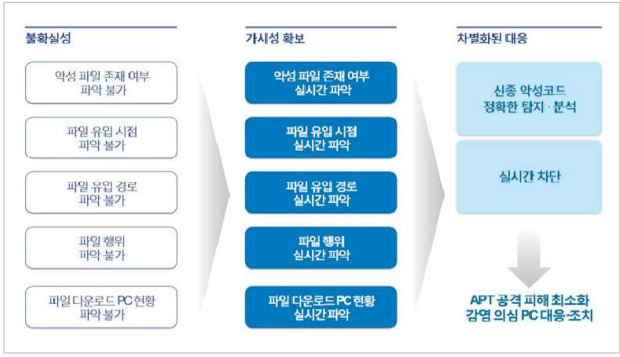
※ ①모바일 악성코드 분석 통계 기능: 사용자의 악성코드 분석 요청 정보를 바탕으로, 전체 요청 수와 분석 알고리즘에 따른 분석 횟수 등 통계 정보 제공 ②모바일 악성코드 분석 이력 기능: 로그인 한 상태에서 사용자의 모바일 악성코드 분석 요청 기록을 조회할 수 있고, 실제 분석 당시 기록을 웹 화면으로 리포팅한 결과를 제공하여 모바일 악성코드 데이터의 위험도를 가시화

[그림] 모바일 악성코드 분석결과 및 시각화 정보



※ 모바일 악성코드 분석 결과 및 시각화 형태 제공: 사용자의 분석 알고리즘 선택 및 APK 파일 업로드 완료 후, 모바일 악성코드 분석 시스템에서 수집한 분석 결과를 바탕으로 웹 화면의 리포트 형태로서 악성코드 위험도, 외부 안티바이러스 탐지 비율, 파일 내부 문자열 정보 등 모바일 악성코드의 위험을 가시화할 수 있는 다양한 정보를 제공

[그림] 데이터바우처 지원 이후 사업 형태



사업의 성과

데이터 바우처로 모바일 악성코드 분석 플랫폼 상용화 개발에 박차

엔에스에이치씨는 최근 수년간 연구해 온 머신러닝 모바일 악성코드 분석 시스템을 상용화하기 위해 그 서비스 모델을 개발하는 중이며, 악성코드 데이터를 정·동적 분석 과정을 거쳐 악성코드 유형 및 통계 데이터 등 데이터 시각화 정보를 대시보드 기능으로 구현하여 웹 표준 기반의 포탈 서비스로 제공합니다. 이렇게 가공된 악성코드 데이터 정보는 JSON 포맷 형태로 저장되는데, 엔에스에이치씨의 분석 포털 서비스에 접속한 악성코드 분석 전문가가 검색, 호출한 악성코드 샘플에 대해 수행된 분석 결과의 수치 및 그래픽 요소를 보고 악성코드의 위험도를 파악할 수 있습니다.

이러한 서비스는 엔에스에이치씨가 개발 중인 '머신러닝 지능형 악성코드 분석 통합 시스템'의 베타테스트 버전으로, 국내외 전시회 또는 자사의 해외 파트너사를 통해 시연 및 교육활동을 진행할 예정입니다.

이를 위해 악성코드 분석, 검색, 유형별 통계 등의 기본적인 기능에다가 우회기술을 적용시켜 지능형 악성코드의 실시간 탐지가 가능하도록 점진적으로 업그레이드를 할 계획입니다.

사업에 대한 의견

특정 데이터를 분석하여 특수 정보 서비스를 창출해 내고자 할 때, 내부에서 보유하지 않는 개발 방법론이나 구현 기법에 대해 보다 효율적으로 업무를 수행하고 비용 절감까지 할 수 있어서 좋았습니다. 데이터 가공기업의 전문성이 보다 다양화되면 중소기업의 사업 아이템 개발에 더욱 실질적인 지원이 가능할 것 같습니다.

1-2. 일반가공

트리즈온

기 업 명	트리즈온
참 여 부 분	일반가공
데이터 활용	장비 및 서버의 보안 이벤트를 통합적으로 분석하여 보안컨설팅 제공
공 급 기 업	엠투어플
성 과	분석 효율성 향상

AI기반 네트워크 침해 정보를 전문적으로 수집하고 분석하다

트리즈온은 빅데이터를 기반으로 네트워크 장비 및 서버의 보안 이벤트를 통합적으로 분석하여 최근 사이버 위협을 탐지, 예측 및 대응하는 SIEM을 제공하고 있으며, 주요제품으로는 비정형데이터 암호화솔루션과 DRM을 대체할 수 있는 SEED암호화솔루션을 보유하고 있습니다. 이를 기반으로 실시간 스트리밍이 가능한 보안 미디어서버 플랫폼을 제공하고 있습니다.

우선, 솔루션 고도화를 위한 개발인력 비용 리스크 부담을 줄일 수 있었고, 이를 계기로 공공기관, 관공서에 대한 네트워크 통합보안관제서비스 영업의 틀을 마련하였습니다. 또한, 고객사의 네트워크, 망분리 구간의 패킷과 웹방화벽 서버 로그를 수집 및 분석하고, 인공지능 기반의 패킷, 로그 분석, 탐지, 통계를 통한 네트워크 통합보안관제서비스를 제공할 수 있게 되었습니다.

[사진] 데이터바우처 사업을 수행하고 있는 모습

사업에 지원하게 된 동기

사이버 보안 시장의 문제점을 해결하기 위하여, 최근의 침해위협의 사전에 탐지 할 수 있을까?

고객의 실제 패킷과 운영 서버들의 이벤트 로그를 통합 분석하여, 통계분석 데이터를 제공하는 전문적인 컨설팅 서비스가 필요하지 않을까?

데이터 활용성의 제한과 가공 데이터의 품질개선을 통하여 신·변종 해킹 기술에 대한 사전탐지와 사전방지 기술을 고도화해야했고, 특히 소기업으로서 실행하기 어려운 신규 사업의 확장 필요성을 위하여 데이터바우처 지원사업이 필요했습니다.

첫째, 보안 분석 통계 리포트를 이용한 컨설팅은 최근 보안관제 서비스 부문에서 중요한 요소로 부각되기 시작했습니다. 특히 시시각각으로 반영되는 침해사고에 대한 컨설팅이 미비한 점이 문제점으로 부각되고 있는데, 이를 위한 최신 해커공격에 대한 테스트 및 이를 탐지하는 데이터셋 제공이 필요했습니다. 둘째, 가공 데이터의 유통 및 활용을 위한 품질, 저작권, 서비스 제공 등에 대한 지원이 필요했습니다. 사이버 해킹에 대한 모의 데이터, 가공 데이터의 생산 활동 및 구매처 정보가 부재한 상황이었고, 전문성 요구로 인해 데이터 가공비용은 높아진 데 비해 가공시장은 미성숙한 상태였습니다. 셋째, 기존 보안으로는 신종 해커에 대한 사전 탐지와 예방 기술 수준이 취약했습니다. 최근 증가하는 신·변종 공격, 대규모 공격, 이상징후에 대한 대응이 어려웠고, 과탐과 오탐이 발생하기도 했습니다.

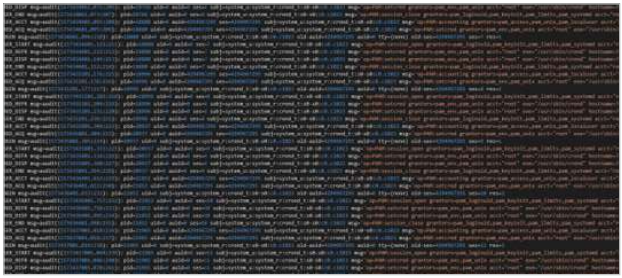
사업 진행과정

PC, 서버 시스템에서 발생하는 비정상 시스템 로그, 패킷 및 트래픽을 분석하여 레포트와 대응 룰셋을 제공하다

우선, 공급기업인 엠투어플 이실드(eSILD) 시스템에서 로그 데이터를 수집하여, 시스템 로그 히스토리 데이터를 산출합니다. 그 후 수집 및 전처리된 데이터를 가공하여 시각화합니다.

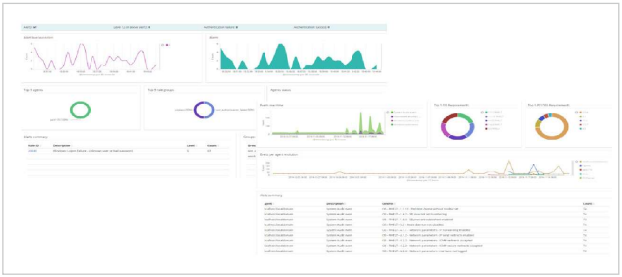
[그림] 시스템 로그 데이터 수집 화면

[그림] 데이터 가공 분석을 통한 가공로그



시스템 로그 데이터와 가공데이터의 처리과정은 다음과 같습니다. 에이전트의 시스템 로그를 실시간으로 취합하고, 원시데이터를 연, 월, 주, 일간으로 가공하고 실시간 에이전트의 행위를 분석합니다. 시스템 로그를 실시간 정보를 분석하여 연간, 월간, 주간, 위험도에 따라 분류하고 시각화하여 통계 분석이 가능하게 됩니다.

[그림] 시스템 로그 데이터와 가공데이터의 시각화



사업의 성과

한정된 시간, 인력, 자원으로 AI 및 머신러닝을 보안관제서비스에 적용하여 분석의 효율성을 향상시키다

이번 데이터바우처 사업을 통해 미탐지 위협의 최소화, 신속한 위협 탐지, 위협 대응시간 단축의 성과를 도출할 수 있었습니다.

향후에는 빅데이터 기반의 분석 플랫폼을 활용하여 기존의 이벤트 중심의 로그분만이 아니라 접속정보, 행위정보, 거래정보, 메타정보 등 다양한 사용자 및 애플리케이션의 로그를 추가로 수집하고, 사용자의 비정상적 이상행위를 탐지하여 내부정보 유출방지, 사기거래탐지, 감사 업무 등에 보다 정확한 보안서비스를 제공할 예정입니다.

사업에 대한 의견

데이터바우처 지원사업은 소기업과 1인 기업처럼 규모가 작은 기업들에게는 업계의 최신 기술동향을 접하고, 실제로 회사의 기술에 접목하거나 한 단계 발전시킬 수 있는 직·간접적으로 아주 좋은 기회입니다. 따라서 더 많은 기업들이 참여할 수 있도록 다양한 홍보 루트를 활성화시키고, 성공 사례가 탄생할 수 있도록 마케팅 및 세일즈에도 전폭적인 지지를 부탁드립니다.

백문불여일견(百聞不如一見)이라는 말이 있듯이, 머리로만 생각하지 말고 실제로 한번 참가해서 몸으로 부딪혀봐야 기술도 회사도 발전합니다!

1-3.
구매

모니터랩

MONITORAPP

기 업 명	모니터랩
참 여 부 분	구매
데이터 활용	URL 정보데이터
공 급 기 업	지란지교소프트
성 과	URL 필터링 및 악성사이트 차당 서비스 제공

해외 URL Category 데이터로 인터넷 접근제어 서비스 경쟁력 강화

모니터랩은 보안소프트웨어 개발 기업으로, 웹방화벽, DB방화벽, 인터넷접근제어 솔루션을 개발하여 공급하며 클라우드를 기반으로 한 보안 서비스를 제공하고 있습니다. 데이터바우처 지원사업을 통해 얻은 URL정보데이터를 모니터랩의 인터넷 접근제어 솔루션에 탑재하여 해외 URL에 대한 Filtering 서비스의 품질을 개선시켰습니다. 또한, 해외 URL Category 데이터를 이용하여 Global URL Filtering 서비스의 품질을 높이고, 내부 개발 자원의 핵심 역량 집중을 실현할 수 있었습니다.

[사진] 모니터랩 사무실

사업에 지원하게 된 동기

Global URL Filtering 서비스 품질을 효율적으로 강화할 방법은 없을까?

인터넷 접근제어 솔루션을 개발하는 기업의 입장에서는 URL Filtering 서비스를 위해 자체적으로 수집·분석하고 있는 국내 URL외에, 해외 URL을 수집하고 분류하여 Global URL Categorization을 통한 악성 사이트 차단과 같은 인터넷 접근제어 서비스의 품질을 높이는 것이 중요합니다. 인터넷접근제어 솔루션의 경쟁력 강화를 위한 해외 URL 데이터의 수집과 분석에 대한 방법론을 고민하고 있던 저희 입장에서 데이터바우처 지원사업은 저희가 필요로 하는 Global URL Category정보를 단기간에 우수한 품질로 구현할 수 있는 좋은 기회라고 생각했습니다. 그래서 본 사업을 통해 지란지교소프트가 제공해주는 데이터를 연동할 수 있는 방법을 모색했고, 성공적으로 데이터를 연동하여 인터넷 접근제어 솔루션의 Global URL Filtering 서비스 품질 경쟁력을 확보할 수 있었습니다.

사업 진행과정

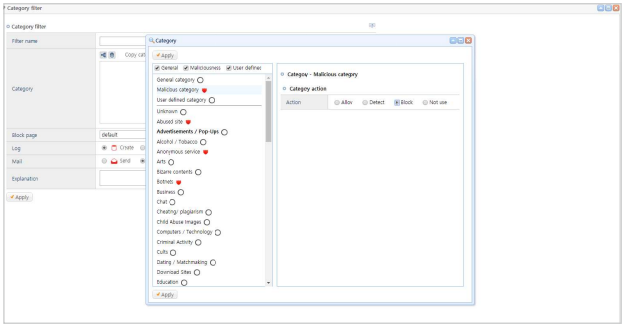
Global URL Categorization 데이터로 악성사이트에 대한 접근 차단율 개선!

모니터랩은 이번 데이터바우처 지원사업을 통해서, URL 정보 데이터를 구매했습니다. 글로벌 기준으로 URL 도메인은 약 100억개 이상이 존재하며 그 중 10억개 이상의 도메인이 활성화되어 운영 중입니다. 이번에 제공받은 URL 정보 데이터는 운영 중인 10억개 이상의 URL 정보를 각 분야별로 분류하여 각 URL에 대한 매칭 정보를 API로 제공함으로써 단기간에 Global URL 분류가 가능하도록 한 데이터입니다. 데이터의 형태는 아래의 이미지와 같습니다.

[그림] 데이터 형태

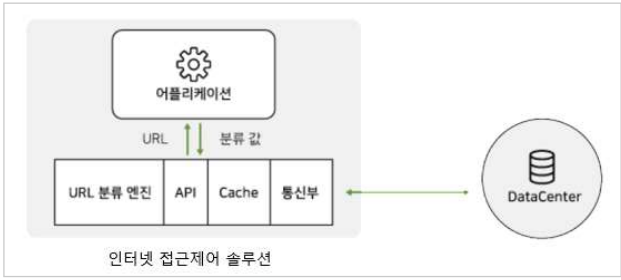
Classification ID	종류	의미	특징	추가 설명
1	Advertisements_and_Pop-Ups	광고	Control	광고컨텐츠제공 및 중계
2	Alcohol_and_Tobacco	술 / 담배	Control	술/담배회사 및 관련 정보
3	Anonymizers	익명	Security	불법 중계(리다이렉트)
4	Arts	예술	일반	국전,미술관등 예술관련
5	Business	비즈니스	일반	기업사이트 / 비즈니스 정보
6	Transportation	자동차(운송수단)	일반	정보/판매/자동차동호회등
7	Chat	채팅	Control	채팅사이트
9	Forums_and_Newsgroups	포럼/뉴스그룹	일반	뉴스공유(개인블로그 포함)
10	Compromised	보안 취약	Security	악성코드삽입/취약 사이트
11	Computers_and_Technology	컴퓨터/기술	일반	정보제공/관련 서비스 사이트
12	Criminal Activity	범죄	Control	범죄모의/범죄정보등
13	Dating_and_Personal	데이트(만남)	Control	결혼정보/데이트주선등
14	Download_Sites	다운로드	Control	SW/동영상등 다운로드
15	Education	교육	일반	교육기관/교육정보

[그림] 데이터 연동 후 솔루션 인터페이스 화면

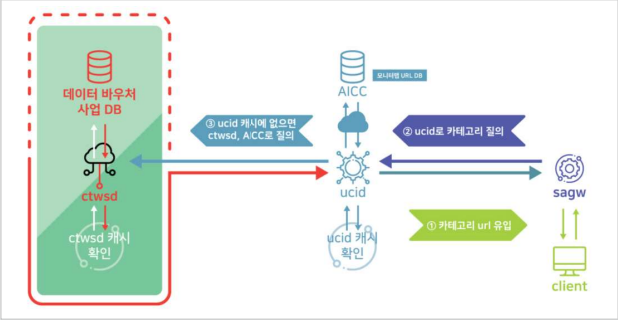


데이터를 인터넷 접근제어 솔루션과 API를 통해 연동하였고, 구매 데이터와 당사 보유 데이터를 캐싱 인터페이스로 연계하여 데이터의 품질을 높였습니다. 또한 미분류 및 오분류, 오탐 데이터에 대한 지속적인 관리로 데이터 라이프 사이클 관리가 가능하게 구연하였습니다.

[그림] 당사 인터넷 접근 제어 솔루션과 API 연동



[그림] 기존 모니터랩 보유 DB와 캐싱 인터페이스를 연계



사업의 성과

데이터바우처로 인터넷 접근제어 서비스의 품질 차별화!

모니터랩은 데이터바우처 지원사업에 참여하여 다양한 사업 성과를 이루어냈습니다. 1.7억 원의 신규 매출을 확보하였을 뿐만 아니라, 해외 악성 URL에 대한 접근 차단 서비스 개발하여 고객사에게 제공할 수 있게 되었습니다. 그 외에도 URL 조회 속도를 80% 개선하였고, 해외 URL의 미분류 및 오분류 데이터의 Feed Back 시스템을 추가할 수 있게 되었습니다. 이러한 성과를 바탕으로 향후 인터넷 접근제어 신사업 매출 확대를 기대하고 있습니다. 향후 모니터랩은 인터넷접근제어 솔루션 매출의 비중이 2019년 2%에서 2022년 16.6%까지 확대될 수 있을 것으로 예상하고 있습니다. 또한 현재 700여개의 기업 가입자를 확보하고 있는 당사 클라우드 기반 SECaaS 서비스에 인터넷 접근 제어 솔루션을 탑재하면, 그 동안 H/W관리의 한계로 관련 솔루션을 도입하기 어려웠던 공공 PC, 학교 실습실 PC 등 공공 시장에 인터넷 접근제어 솔루션을 서비스 형태로 공급이 가능하기 때문에 서비스 이용자를 더 많이 확보할 수 있을 거라고 기대하고 있습니다.

사업에 대한 의견

데이터바우처 지원사업은 중소기업 입장에서 다양한 데이터를 기반으로 하는 제품 및 서비스를 개발할 때 핵심역량에 집중할 수 있도록 도움을 주는 효과적인 지원 사업이라고 생각합니다. 향후 데이터바우처 지원사업이 보다 활발한 관련 전후방 산업의 마중물 역할을 하기 위해서는 데이터 공급기업의 정보가 사전에 충분히 공유되어 수요기업이 보다 쉽게 데이터에 접근할 수 있도록 해야 한다고 생각합니다. 또한 정보 공유 방안과 데이터 연동기간을 고려할 때 2~3년간 정도의 다년 사업으로 지원되어, 더욱 효과성을 낼 수 있도록 사업 기간에 대한 고려가 필요하다고 생각합니다.

데이터바우처 사업은 중소기업 입장에서 아주 매력적인 지원 사업입니다. 회사가 개발하고 있는 제품 및 서비스에 필요한 데이터가 있다면, 직접 확보함과 함께 데이터바우처 사업을 통해 필요한 데이터를 발굴하고 연동하여 가치를 확인하는 방안을 동시에 모색할 것을 적극 추천합니다. 다만 필요한 데이터를 발굴하고 가공이나 연동을 하기 위한 기술적인 검토를 위해 사전에 충분한 조사를 진행할 것을 조언하고 싶습니다.

1-3.
구매

진앤현시큐리티

기 업 명	진앤현시큐리티
참 여 부 분	구매
데이터 활용	위협 정보 데이터 구매
공 급 기 업	엔피코어
성 과	자사 플랫폼 컨텐츠 품질 향상

기존 보안플랫폼 시스템 사업 및 보안운영관제 서비스 질 향상으로 사업의 안정성을 높이다.

진앤현시큐리티는 지난 2000년 설립 이래 보안시스템 구축 및 운영을 통해 체득한 보안 관리경험을 바탕으로 보안관리 전문기업을 목표로 고객 맞춤형 보안 관리서비스를 체계적으로 제공해왔습니다.

보안관리 전문기업인 진앤현시큐리티는 고객의 요구와 시장변화를 신속하고 정확하게 파악하여 최신 기술력을 바탕으로 보안 관리자의 경험을 나누기 위해 노력해 왔습니다.

최근에는 정보보안 패러다임의 변화와 고도화되고 있는 보안 위협에 대비하여 보안 관리자들의 효율적인 보안관리 전략수립과 보안 운영을 위해 최선을 다하고 있습니다. 이번 데이터바우처 지원사업을 통해서는 보안위협정보 데이터를 구매하였으며, 이를 통해 자사의 플랫폼 서비스 품질을 향상하고, 고객사에 데이터를 부가서비스 형태로 제공할 수 있게 되었습니다.

[사진] 진앤현시큐리티 사무실

사업에 지원하게 된 동기

중소기업용 SECUMOM에 보안정보를 제공함으로써 안정적인 보안활동을 지원할 수 있을까?

중소기업용 SECUMOM에 보안정보를 제공함으로써 안정적인 보안활동을 지원할 수 있을까?

최근 점점 더 교묘해지는 보안위협이 지금까지와는 비교할 수 없을 만큼 광범위하고 치명적이 되어가고 있는 상황에서, 기업의 정보보안 시스템 구축은 날이 갈수록 그 중요성이 더해지고 있습니다. 하지만 많은 기업에서 정보보안을 위한 활동을 하고 있지만 사전대응에 대한 정보는 많이 부족한 상황이며, 정보보안솔루션 운영을 중심으로 사이버해킹에 대비하는 정도입니다. 기업에서 자체적으로 사전에 대비할 수 있는 사이버위협정보를 제공받기도 어려울 뿐만 아니라, 시스템구축 및 전문기업의 서비스를 제공받기 위해서는 많은 비용과 전문 인력이 필요합니다. 그리고 특히, 자본과 인력이 부족한 중소기업의 경우 정보보안의 사각지대에 놓이게 되는 경우가 많습니다.

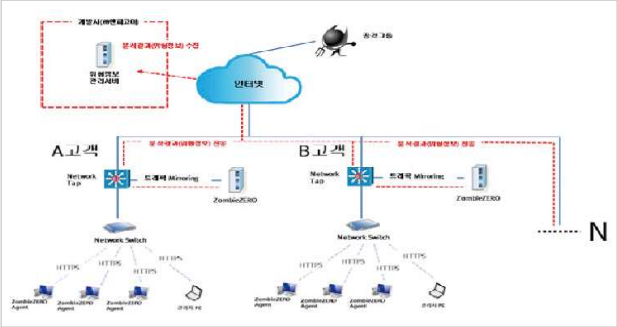
자사의 차세대보안관리플랫폼인 SEMUMOM은 중소기업용 플랫폼으로서 보안위협정보를 제공하여 중소기업의 안정적인 보안활동을 지원하고자 하는 목적에서 기획된 솔루션입니다. SECUMOM은 솔루션 업체 등에서 제공하는 보안위협 정보 외에는 정보를 확인하기 어려운 점을 해소하기 위해 탄생하였습니다. 데이터바우처 지원사업에는 SECUMOM의 서비스 품질을 향상시키기 위해 참여하게 되었습니다. 데이터바우처 지원사업을 통해 공급기업의 사이버 위협정보 데이터를 구매하였고, 이를 활용하여 자사의 SECUMOM 서비스에 사이버 위협정보를 업로드하여 고객사에게 양질의 보안 서비스를 제공할 수 있게 되었습니다.

사업 진행과정

공급기업의 보안위협 데이터를 가공하여 SECUMOM을 통해 고객사에 제공하다

공급기업에게 제공받은 보안위협데이터를 활용하여, 자사의 차세대보안운영관리플랫폼인 SECUMOM에 연동하는 과정을 통해 서비스를 상용화할 예정입니다. 공급기업의 위협데이터는 행위기반 솔루션에서 1차로 수집된 정보를, 공급기업 사이버 안전센터에서 2차로 수집 및 가공한 데이터입니다. 데이터 수집 과정은 아래 그림과 같습니다.

[그림] 위협데이터 수집과정



수집한 데이터를 자사 플랫폼에 웹 다운로드 방식으로 업데이트하여, 고객사에게 서비스를 제공할 예정입니다. 플랫폼의 실제 모습은 아래 사진과 같습니다.

[그림] 플랫폼 캡처 화면

사업의 성과

데이터바우처로 고객사에 더 양질의 보안위협정보를 제공하다

공급기업인 엔피코어의 사이버위협정보 포털을 통해 만들어진 최신 위협정보들을 자사의 플랫폼에 적용해, 더 고도화된 차세대보안운영관리플랫폼을 상용화할 예정입니다.
또한 이러한 최신위협정보들을 현재 자사가 보유한 고객사와 연동 또는 공급을 통해 부가서비스 형태로 서비스를 개시할 예정입니다.

이러한 제품 및 서비스를 자사의 주력제품으로 활용하여, 신규 사업 입찰 시 입찰 경쟁력을 확보할 계획입니다. 데이터바우처 지원사업을 통해, 자사의 제품 경쟁력 향상과 사업 수주율 제고를 기대하고 있습니다.

사업에 대한 의견

데이터바우처 지원사업에 수요기업으로 참여하면서, 사업에 꼭 필요한 형태의 데이터를 제공받는다는 것이 생각보다 쉽지 않다는 생각이 들었습니다. 각각의 수요기업이 필요한 맞춤형 데이터를 제공받기 위해서는 더 다양한 형태의 데이터와 서비스가 제공되어야 할 것 같습니다. 또한 이러한 양질의 데이터 지원이 가능하려면 사업 예산이 확대되어야 하지 않을까하는 생각이 듭니다.

반면에 데이터바우처 지원사업은 정보보호 업계에 좋은 기회가 될 수 있다고 생각합니다. 저희와 같은 보안전문 기업들은 지금도 끊임없이 진화하고 있는 각종 신변종 악성코드로부터 안전한 인터넷 환경을 조성하기 위해 존재합니다. 악성코드가 끊임없이 변화하고 있는 만큼, 저희 또한 발 빠르게 최신 위협정보들을 보유하고 있어야만 최상의 대응이 가능할 것입니다. 각 기업에 흩어져있는 최신 위협정보데이터를 데이터바우처 지원사업을 통해 공유한다면, 국내 보안서비스 제품의 품질 향상에 큰 도움이 될 수 있을 것이라고 생각합니다. 신규 정보보호 분야의 기술을 선도하고, 정보보호 산업 육성 및 기술 선진국으로서의 위상을 확립하려면 이러한 데이터 공유의 과정이 꼭 필요합니다. 많은 보안기업들이 수요 및 공급기업으로 참여한다면, 미래 국내 정보보호 산업의 발전에 큰 도움이 될 수 있을 것입니다. 많은 기업들의 적극적인 참여를 부탁드리고 싶습니다.

2-1.
AI가공

내프터

now and after
NAFTER

기 업 명	내프터
참 여 부 분	AI 가공
데이터 활용	심박 등 생체 정보 데이터 AI 학습셋 가공
공 급 기 업	엘렉시
성 과	생체 정보 기반 응급 알림 서비스 고도화

생체 데이터 분석으로 건강하고 행복한 사회복지 실현에 기여하는 건강 케어 솔루션의 사업성을 높이다

내프터는 클라우드 및 인공지능 ICT기업으로, 클라우드 컴퓨팅과 망분리 시스템의 조달 및 양산을 주 사업으로 진행하고 있으며, 노인을 포함한 고령화 사회 복지에 대비하는 솔루션으로 인공지능 기반 생체 정보 데이터 분석 및 건강 케어 솔루션 개발을 진행하고 있습니다. 그 외 시스템 반도체 관련 기획 및 지원 사업 등을 진행하고 있습니다. 데이터바우처 AI가공 분야에 수요기업으로 참여하였으며, 데이터바우처 지원사업을 통해 제공받은 데이터 가공 서비스로 기업의 신성장 동력을 확보하는 밑거름이 되는데 큰 도움을 받았습니다.

[사진] 내프터 대표님

사업에 지원하게 된 동기

빅데이터와 AI를 이용하여, 생체 정보의 이상 징후를 보다 정교하게 탐지할 수 있을까?

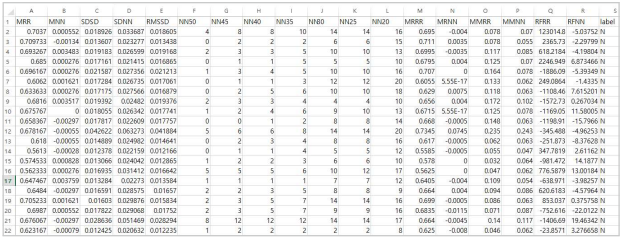
심박, 호흡 등의 사람 생체 정보를 이용하여 건강 이상상태를 탐지하는 서비스를 개발 중인 내프터는, 더 정확한 이상 징후 탐지를 위한 '다각도의 심박변이도 특성 분석 및 DB'의 축적 및 분석 서비스가 필요하였습니다. 사람의 심박은 다양한 생체 신호를 담고 있으며, 이를 분석함으로써 향후 다양한 병증과 이상 징후를 예측해 낼 수 있는 정보라고 할 수 있습니다. 특히, 독거노인이나 복지 사각지대에 놓여있는 대상자들을 위한 응급알림서비스 등에 적용할 경우, 사회문제화 되고 있는 고독사 등의 문제를 경감시킬 수 있는 일환으로 활용될 수 있는데요, 실제 보건복지부 담당 관계자도 이런 필요성을 강하게 요청받고 있다고 합니다. 기존의 서비스는 대상자의 생체 신호 분석보다는 화재, 가스 누출과 같은 주변 환경 위급상황에 대해서만 단순히 탐지하는 수준입니다. 하지만 심박과 같은 대상자의 생체 정보의 이상 징후를 탐지 및 예측할 수 있다면, 서비스의 질과 기회비용을 줄일 수 있다고 할 수 있습니다. 따라서 데이터바우처 지원사업을 통해 생체 정보를 다각도로 분석 및 DB화하고, 보다 정확도 있는 AI가공을 할 수 있다면, 서비스 고도화와 BM 확대를 위해 매우 유용한 기회라고 생각했습니다. 그래서 적극적으로 데이터바우처 지원사업에 지원을 하게 되었습니다.

사업 진행과정

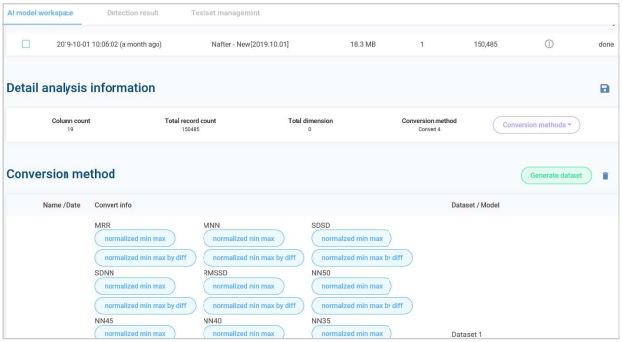
다각화된 심박 등의 생체 데이터 분석과 보다 정확도 있는 이상 징후 탐지 및 예측에 기반 한 응급알림서비스 고도화 서비스

기존에 자사에서 보유 중이던 '심박 등 생체 정보 데이터'가 AI학습셋에 적용될 수 있도록 다양한 각도로 가공하는 과정을 거쳐 AI관제 시스템에 탑재했습니다. 자사는 오픈 DB를 기초로, 엑셀에 정리하고 다양한 각도의 심박 변이와 관련된 파라미터를 생성하여 다량의 데이터 가공 및 분석 작업을 진행했습니다.

[그림] 대상자 순서, 심박변이 파리미터 데이터 가공(A~R), 이상 징후 라벨(S)

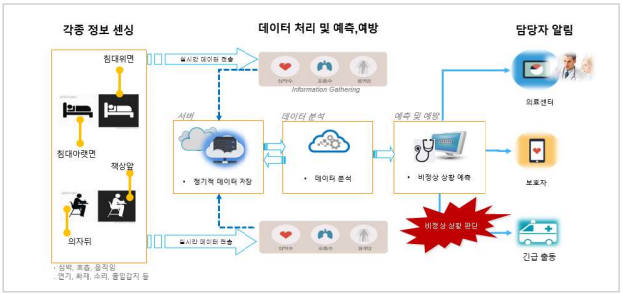


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	MRR	MNN	SDSD	SDNN	RMSSD	NN50	NN45	NN40	NN35	NN80	NN25	NN20	MRRR	MRNN	MMRR	MMNN	RFRR	RFNN	label
2	0.7037	0.000552	0.018926	0.033687	0.018605	4	8	8	10	14	14	16	0.695	-0.004	0.078	0.07	1230148	-5.03752	N
3	0.709733	-0.00134	0.013607	0.023277	0.013438	0	2	2	2	6	6	15	0.7111	0.0035	0.078	0.055	2365.73	-2.29799	N
4	0.693267	0.003483	0.019183	0.026599	0.019168	2	3	3	5	10	10	13	0.6995	-0.0035	0.117	0.085	618.2184	-4.19804	N
5	0.685	0.000276	0.017161	0.021415	0.016865	0	1	1	5	5	5	10	0.6795	0.004	0.125	0.07	2246.949	6.873466	N
6	0.696167	0.000276	0.021587	0.027356	0.021213	1	3	4	5	10	10	16	0.707	0	0.164	0.078	-1886.09	-5.39349	N
7	0.6962	0.001621	0.017284	0.026735	0.017061	0	1	1	3	12	12	20	0.6055	5.55E-17	0.133	0.062	2491.864	-1.14335	N
8	0.633633	0.000276	0.017175	0.027566	0.016879	0	2	5	6	10	10	18	0.629	0.0075	0.118	0.063	-1108.46	7.615201	N
9	0.6816	0.003517	0.019392	0.02482	0.019376	2	3	3	4	4	4	10	0.656	0.004	0.172	0.102	-1572.73	0.267034	N
10	0.675767	0	0.018055	0.026342	0.017741	1	2	4	6	9	10	13	0.6715	5.55E-17	0.125	0.078	-1169.05	11.58005	N
11	0.658367	-0.00297	0.017817	0.022609	0.017757	0	0	1	2	8	8	14	0.668	-0.0005	0.148	0.063	-1198.91	-15.7966	N
12	0.678167	-0.00055	0.042622	0.063273	0.041884	5	6	6	8	14	14	20	0.7345	0.0745	0.235	0.243	-345.488	-4.96253	N
13	0.618	-0.00055	0.014889	0.024982	0.014641	0	2	3	4	8	8	16	0.617	-0.0005	0.062	0.063	-251.873	-8.37628	N
14	0.5613	-0.00028	0.012378	0.022159	0.012166	0	1	1	4	5	5	12	0.5585	-0.0005	0.055	0.047	347.7819	2.61162	N
15	0.574533	0.000828	0.013066	0.024042	0.012865	1	2	2	3	6	6	10	0.578	0	0.032	0.064	-981.472	14.1877	N
16	0.562333	0.000276	0.016935	0.031412	0.016642	5	5	5	6	10	12	17	0.5625	0	0.047	0.062	776.5879	13.00184	N
17	0.647467	0.003759	0.013284	0.02273	0.013084	1	1	1	1	7	7	12	0.6405	-0.004	0.109	0.054	-638.971	-3.98257	N
18	0.6484	-0.00297	0.016591	0.028575	0.01657	2	2	3	5	8	8	9	0.664	0.004	0.094	0.086	620.6183	-4.57964	N
19	0.705233	0.001621	0.01603	0.029876	0.015834	2	3	5	7	14	14	16	0.699	-0.0005	0.086	0.063	853.037	0.375758	N
20	0.6987	0.000552	0.017822	0.029068	0.01752	2	3	5	7	9	9	16	0.6835	-0.0115	0.071	0.087	-752.616	-22.0122	N
21	0.676067	-0.00297	0.028636	0.051469	0.028294	8	12	12	12	14	14	17	0.664	-0.0045	0.14	0.117	-1406.69	19.46342	N
22	0.623167	-0.00079	0.012425	0.020632	0.012235	1	2	2	2	2	2	8	0.625	-0.008	0.046	0.062	-23.8571	3.276658	N



※ 학습 모델 #4의 경우, 150,485개의 Normal 데이터가 활용되었으며, 각 파라미터 데이터는 전처리 과정을 거친 후, 학습 데이터로 이용되었음.

[그림] 데이터 활용 방법에 대한 개념도



비접촉 생체 센서로부터 추출된 데이터 중, 가장 중요하다고 판단되는 심박에 대한 데이터 신뢰도를 높이기 위해 심전도(ECG) 데이터를 심박변이도(HRV)로 변경하였습니다. 또한 인공지능 학습 모델의 안정성과 정확도를 높이기 위해 장·단기 심박변이도 특성을 나타내는 18종의 파라미터로 데이터를 가공하여 학습에 필요한 데이터를 충분히 확보하도록 하였습니다. 데이터의 양, 심박변이 파라미터 들 간의 상관관계, 전처리 수준 등의 상관관계를 분석하기 위해 다양한 조합으로 테스트가 진행되었으며, 최적의 학습율을 탐지하기 위해 학습과 테스트를 반복 시행하였습니다.

사업의 성과

데이터바우처로 대상자의 생체 정보 기반 응급알림서비스 고도화에 한 걸음 다가가다

생체 데이터에 대한 다각도의 분석 방법이 도출되었으며, 그에 따른 데이터 가공을 통해 인공지능 학습 및 테스트에 맞는 빅데이터를 확보할 수 있게 되었습니다. 또한, 데이터 특성에 맞는 전·후처리 기법을 도입하여 보다 정확도 있게 학습될 수 있는 플랫폼을 구축하게 됨에 따라 향후 사업화 진행시, 보다 경쟁력 있는 플랫폼과 서비스를 제공할 수 있을 것으로 예측됩니다. 또한, 데이터바우처가 계기가 되어 국내 유수 의료기관과 MOU를 맺게 되어, 향후 국내에 특화된 데이터에 기반하여 연구를 진행할 수 있는 계기도 마련되었습니다.

2021년을 목표로, 서비스 출시를 위해 보다 정교하고 고도화된 플랫폼 완성을 진행할 예정입니다. 또한, 국내 많은 전문 의료 전문가들이 지적하는 해외 데이터 기반의 인공지능 분석에 따른 문제점에 동의하는 입장으로서 국내 데이터 확보와 가공을 통해 한국 사람들의 데이터에 특화된 서비스가 가능하도록 인공지능 플랫폼을 개선해 나갈 예정입니다. 또한, 궁극적으로는 Edge 인공지능을 구현하기 위해 데이터바우처 사업 결과를 활용할 예정이며, 이는 Edge AI를 완성하기 위한 훌륭한 벤치마킹 데이터가 될 것으로 확신합니다. 그리고 현장에서 이용 가능한 서비스가 될 수 있도록 지자체 및 의료 기관 협력을 강화하여 실증 사업 강화에 좀 더 매진할 계획입니다.

사업에 대한 의견

데이터 가공 기업, 플랫폼 공급 기업, 서비스 기업 등이 데이터바우처 사업 공간에서 만나서 시너지 효과를 낼 수 있었던 좋은 기회였습니다. 특히, 초기에 기업이 부딪히는 인력, 자금, 기술 확보 문제에 직면했을 때 중소기업들 각사가 가지고 있는 강점과 장점을 살려서 하나의 목표를 향해서 기술 및 서비스 개발을 진행하는 것을 통해 국내 산업 발전의 모멘텀이 강화될 것으로 생각합니다. 앞으로는, 보다 다양한 분야의 기업들이 데이터바우처 사업을 통해 협력하고 각사가 가지고 있는 강점들을 모아서 시너지를 낼 수 있는 다양한 기회가 만들어지기를 희망합니다.

앞으로의 산업 생태계는 한 기업이 모든 것을 담당하고, 개발하는 시대에서 서로 협력하고, 이종 간의 산업 간에도 콜라보레이션이 가능해야 하는 시대로 넘어가고 있습니다. 따라서 현재의 서비스 시스템을 보다 지능화 및 고도화 하고자 하는 기업들은 데이터바우처 사업에 적극적으로 참가하여 다양한 기업들과 협력할 수 있는 기회를 만들기를 추천드립니다.

2-2.
일반가공

엘컴텍

기 업 명	엘컴텍
참 여 부 분	일반가공
데이터 활용	상황 혹은 객체 인식 대상 이미지 데이터 정제 가공
공 급 기 업	셀렉트스타
성 과	홈 내의 정 객체, 동 객체, 상황 데이터 셋을 활용하여 집 내부의 물체 및 상황인식 서비스 상시 품평 및 시범서비스 완비

청소·요리·취침 등 고수준 상황을 파악하여 더 스마트한 홈을 만들다.

엘컴텍은 스마트홈, 무인기 등에 필요한 지능형 상황 인지 서비스 개발을 주 사업으로 하고 있습니다. 국내의 가구 구조 및 구성물에 특화된 데이터 셋을 더 확보하기 위하여 이번 데이터바우처 지원사업에 참여를 하였습니다. 이번 사업에서는 스마트 홈 내에서 이루어지는 상황을 조사, 분석하여 지능형 홈서비스를 제공하기 위한 '스마트미러' 개발을 진행하였습니다. 만족스러운 결과를 갖게 되어 기쁩니다.

[사진] 스마트미러

사업에 지원하게 된 동기

데이터를 활용하면 상황을 인지하는 기술의 정확도를 높일 수 있을까?

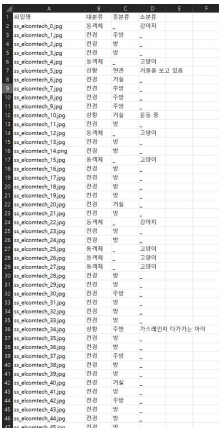
집 내부의 상황을 인식하는 서비스를 개발 중인데 보다 더 높은 정확도와 인식범위를 넓히기 위하여 추가 객체 및 상황 서술 데이터가 필요했습니다. 상황인지 서비스 개발에 있어서 가장 큰 문제점은 학습할 데이터가 부족하다는 점이었습니다. 바로 집 안에서 일어나는 상황들에 대한 데이터를 수집해야했기 때문입니다. 데이터가 많이 없을 뿐만 아니라 개인정보 문제가 있을 수 있기 때문에 풀기 어려운 숙제였습니다. 그래서 데이터 바우처 지원사업에 지원하였고, 실제로 필요로 하는 데이터를 수집하여 현재 고수준 영상 상황 인지 기술 개발을 훨씬 원활하게 진행하고 있습니다.

사업 진행과정

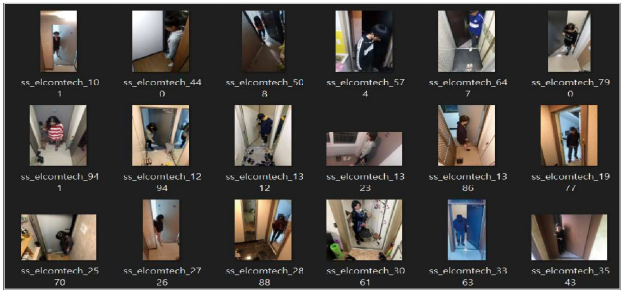
이미지데이터, 이러한 활용법이?

홈 내에 있는 객체 및 상황 학습을 위한 정 객체, 동 객체, 상황들에 대한 데이터를 수집하여 학습데이터로 가공합니다.

[그림] 태깅 정보 종합 csv파일

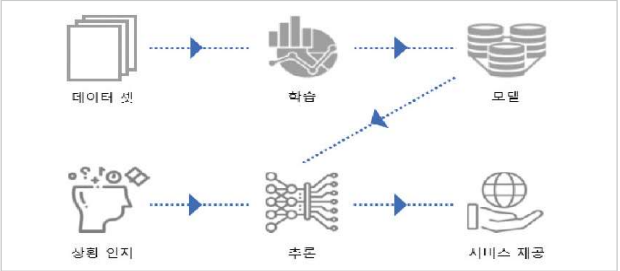


[사진] 수집 완료 데이터 예시 (사람이 현관으로 들어오는 상황)



수집되고 정제된 이미지 데이터를 인공지능기반의 객체인식 및 상황인식 학습네트워크로 학습을 지행하여 추론용 네트워크를 확보하였습니다. 확보된 추론용 네트워크를 활용하여 홈 내에서 수집된 영상이미지를 분석하여 이미지 내의 객체와 상황을 추론하였습니다. 상황이 추론이 되면 그에 맞게 설계된 서비스를 자동으로 시작하게 됩니다.

[그림] 수집 데이터 활용도

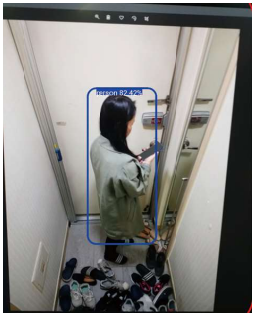


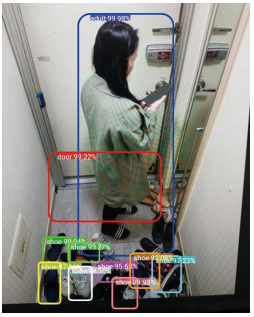
사업의 성과

데이터바우처로 성공적인 딥러닝 영상상황인지 기술 확보!

기존의 객체상황인지 기술은 사물을 인식하지 못하고 사람만 인식했습니다. 데이터 활용 후에는, 개발한 기술로 테스트를 해보았을 때 수집한 데이터들을 모두 정상적으로 인식하는 모습을 확인할 수 있었습니다.

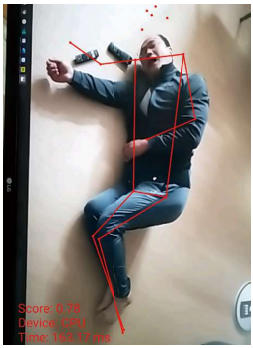
[사진] 현관의 모습 (사람, 사물)





기존의 기술은 사람이 쓰러진 사진으로 비교를 해보았을 때 뒷모습이거나, 옆으로 누워 있는 경우 잘 인식을 못했습니다. 데이터 활용 후에는 사용자의 상태 및 행동을 분석할 수 있게 되어, 사람 포즈 기반의 기술로서 쓰러짐 등과 같은 위험 상태를 탐지하여 위험 상황에 대처할 수 있는 서비스를 제공할 수 있게 되었습니다.

[사진] 쓰러져 있는 사람



데이터바우처 지원사업을 통해 개선된 객체 상황 인지 프로그램은 종래 객체 상황인지 프로그램 대비 확연히 높아진 수준의 객체 인식률을 확인할 수 있었습니다. 현재 자사의 고객사에 영상 상황인지 기술을 제공하여 시제품 평가 및 체험단 운영을 통해 상용서비스 준비를 할 예정입니다. 향후 기술의 안정화 및 구체화 된 서비스 확보를 통해 2021년에 출시할 예정입니다.

사업에 대한 의견

기업마다 사업규모를 차등화하고 규모에 맞도록 수행기관을 차별화 하는 것이 필요하다고 생각합니다.

그리고 만약 사업에 지원하게 된다면 적합한 형태의 데이터 수집 및 가공방법에 대한 사전 준비가 필요합니다. 또한, 수집과 가공 과정 중에 수많은 시행착오를 갖게 되며 제한된 사업비 내에서 결과가 충족되지 않을 수도 있습니다. 따라서 사업에 참여하기 전 가능한 적은 규모라도 학습데이터를 수집 및 가공하여 적용해 보는 것이 필요합니다.

2-3.
구매

피타그래프

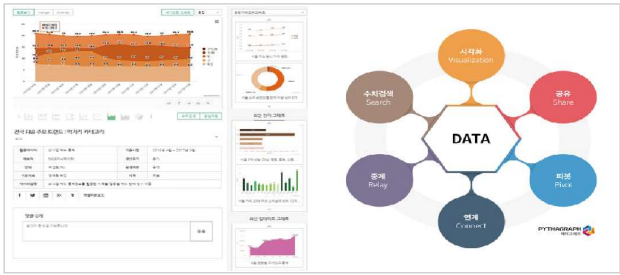
다양한 데이터 확보를 통해
시각화된 그래프 데이터 서비스를 제공하다.

[사진] 피타그래프 사무실

피타그래프는 포털 및 기타 인터넷 정보매개 서비스업체로서 숫자기반데이터의 지능형 동적시각화(IDV: Intelligence Dynamic Visualization)기술을 기반으로 한 그래프 데이터 플랫폼 서비스인 피타그래프닷컴(www.pythagraph.com)을 운영 하고 있습니다. IDV기술은 동일한 차원을 기준으로 데이터가 중첩·병합되거나, 그래프 변형과 드릴링이 자유롭게 되도록 하는 새로운 개념의 검색엔진 및 시각화 기술입니다.

주요제품으로 엔진형 솔루션인 피타그래프-레드, 대쉬보드인 피타그래프-대쉬 솔루션과 데이터 중계 엔진인 OJAK을 보유하고 있습니다. 데이터바우처 지원사업을 통해서 고품질의 콘텐츠를 사용자에게 제공할 수 있어서 사업에 도움이 되었습니다. 고품질의 유료 데이터를 확보하고, 고객에게 시각화된 그래프 콘텐츠 제공을 통해 서비스 이용률 확대에 기여할 수 있었습니다.

[그림] 피타그래프닷컴 홈페이지 화면



사업에 지원하게 된 동기

세상의 모든 숫자를 모아 서로 연결하는 것이 가능할까?

유료로 유통되는 고품질의 데이터는 많은 비용이 소요되어 데이터를 확보하는 데 어려움이 있었고, 그래서 데이터바우처 사업을 통해 지원받고자 하였습니다. 피타그래프는 숫자 기반의 데이터 플랫폼 비즈니스 모델을 가지고 있습니다. 숫자 데이터는 개인의 몸무게, 신용도, 기업의 매출, 주가, 음식, 국가의 인구, GNP, 기온이나 강우량, 행성의 질량이나 원자의 크기 등 무한하고 다양하지만, 종이책, html, 사진, 표 등의 형태로 흩어져 있어 활용도가 낮습니다. 피타그래프는 이러한 숫자 데이터를 시각화하고 동적으로 제공해주며, 그래프 간 병합/중첩, 차원 및 계층 변경, 수치검색(상관도) 등 다양한 기능의 제공을 통해서 활용도를 극대화하기 위해 데이터바우처 지원사업에 지원하게 되었습니다.

사업 진행과정

서울, 부산, 인천 지역의 외식업 데이터를 시각화된 그래프로 제공

외식업 관련해서, 20여종의 메뉴별 결제단가 등의 데이터를 제공받았습니다. 이 데이터는 서울, 부산, 인천 지역의 광역시도, 시군구, 읍면동 데이터 셋으로 구성되어 있었으며, 이 데이터 셋을 변환 및 통합하여, 데이터 연계를 통해 시각화된 그래프를 서비스에 적용하였습니다.

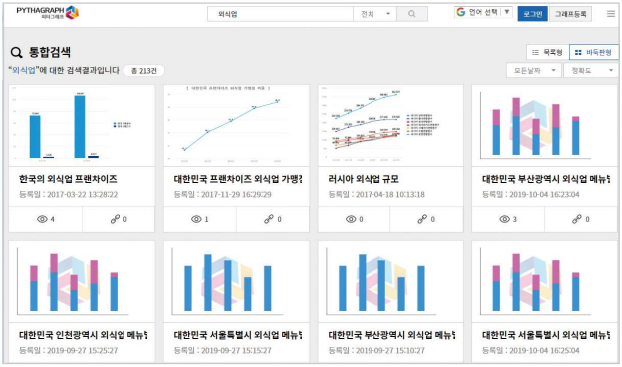
[그림] 광역시도 시군구 메뉴별 결제단가 및 결제건수 비중

기준지역단위	기준년월	년도	월	국가	광역시도	시군구	읍면동	메뉴명	결제건수비중
광역시도	201808	2018	08	대한민국	인천광역시			국산맥주	7.6
광역시도	201809	2018	09	대한민국	인천광역시			국산맥주	7.3
광역시도	201810	2018	10	대한민국	인천광역시			국산맥주	6.5
광역시도	201811	2018	11	대한민국	인천광역시			국산맥주	6.3
광역시도	기준지역단위	기준년월	년도	월	국가	광역시도	시군구	읍면동	메뉴명	결제건수비중
광역시도	시군구	201808	2018	08	대한민국	인천광역시	강화군		국산맥주	3.1
광역시도	시군구	201809	2018	09	대한민국	인천광역시	강화군		국산맥주	2.9
광역시도	시군구	201810	2018	10	대한민국	인천광역시	강화군		국산맥주	2.3
광역시도	시군구	201811	2018	11	대한민국	인천광역시	강화군		국산맥주	2.4
광역시도	시군구	201812	2018	12	대한민국	인천광역시	강화군		국산맥주	2
광역시도	시군구	201901	2019	01	대한민국	인천광역시	강화군		국산맥주	2.1
광역시도	시군구	201902	2019	02	대한민국	인천광역시	강화군		국산맥주	1.9
광역시도	시군구	201903	2019	03	대한민국	인천광역시	강화군		국산맥주	1.7
광역시도	시군구	201904	2019	04	대한민국	인천광역시	강화군		국산맥주	1.9
광역시도	시군구	201905	2019	05	대한민국	인천광역시	강화군		국산맥주	1.9
광역시도	시군구	201906	2019	06	대한민국	인천광역시	강화군		국산맥주	1.9
광역시도	시군구	201907	2019	07	대한민국	인천광역시	강화군		국산맥주	1.9
광역시도	시군구	201808	2018	08	대한민국	인천광역시	강화군		국수	1.6
광역시도	시군구	201809	2018	09	대한민국	인천광역시	강화군		국수	2
광역시도	시군구	201810	2018	10	대한민국	인천광역시	강화군		국수	1.8
광역시도	시군구	201811	2018	11	대한민국	인천광역시	강화군		국수	1.1
광역시도	시군구	201812	2018	12	대한민국	인천광역시	강화군		국수	1.9
광역시도	시군구	201901	2019	01	대한민국	인천광역시	강화군		국수	1.8
광역시도	시군구	201902	2019	02	대한민국	인천광역시	강화군		국수	2
광역시도	시군구	201903	2019	03	대한민국	인천광역시	강화군		국수	1.9
	시군구	201904	2019	04	대한민국	인천광역시	강화군		국수	1.9
	시군구	201905	2019	05	대한민국	인천광역시	강화군		국수	1.9
	시군구	201906	2019	06	대한민국	인천광역시	강화군		국수	4.1
	시군구	201907	2019	07	대한민국	인천광역시	강화군		국수	3.6

사업의 성과

서울, 부산, 인천 지역의 외식업 관련 그래프가 약 70배 증가하다!

[그림] 시각화된 상세 그래프 이미지



기존에는 외식업 시각화 그래프를 3건만 제공하였으나, 데이터 활용 이후에는 외식업 관련 시각화 그래프가 210건 추가되어, 총 213건의 그래프를 제공하고 있습니다.

다양한 콘텐츠의 지속적인 확보 및 사용자 간의 데이터 공유를 통해 이용률을 증대했습니다. 이번 사업을 통해 제공받은 숫자 데이터를 글로벌 비즈니스에 진출하는 기반으로 삼고자 합니다. 2019년에는 서비스 기능에 대한 고도화(수치검색 등)를 통해서 서비스 품질이 향상되었다면, 향후에는 다국어 번역 기능 구현을 통해 글로벌 비즈니스 모델을 발굴할 예정입니다.

사업에 대한 의견

데이터바우처 지원사업은 데이터 기반의 사업이므로, 서비스 제공이 지속적으로 연계가 되어야 한다고 생각합니다. 또한 아쉬운 점이 있다면, 데이터 구매 비용이 다소 적어 다양한 데이터를 활용하는 데에는 한계가 있었습니다.

향후 데이터바우처 지원사업에 참여하는 기업은 데이터 구매 시 활용 목적에 맞는 데이터를 보유한 공급기업과 지속적으로 사업을 수행할 수 있도록 초기단계부터 면밀한 협의가 필요하다고 생각합니다.

03.
기타

범고랩(AI가공)

활용서비스	인카엔터테인먼트 컨텐츠 개발
주요내용	자동차 이미지 데이터 300,000건을 가공하여, 인카엔터테인먼트 컨텐츠 제작을 위한 이미지 가공 tool을 개발 및 활용
기대효과	성공적인 개발 시 자율주행자동차 시장의 성장과 맞물린 IN-CAR-ENTERTAINMENT 시장의 확장에 대비할 수 있는 미래 자동차 시장의 필수적인 기술로서 산업 발전에 기여할 수 있는 것으로 예상
활용성과	자동차의 라벨링이 가능하게끔 하는 소프트웨어 툴 개발 및 현대 자동차 이노베이션 센터 제로원에 참여하여 제품 시연

[사진] 자동차 라벨링 소프트웨어 툴(예시)

[사진] 현대 자동차 이노베이션 센터 제로원 시연

시옷(일반가공)

활용서비스	IoT 디바이스 취약점 DB화 및 시각화분석
주 요 내 용	무선 네트워크 취약점 관련 데이터를 분석 및 가공하여 IoT 디바이스에 대한 정상/비정상 예측 및 취약점 점수화까지 가능한 클라우드 기반 통합 관리 서비스 개발
기 대 효 과	보안데이터 분석을 정규 예측모델(정상/비정상, 위험 등급화) 개발 및 사업화 비용 절감
활 용 성 과	멀티 네트워크 기반 IoT 디바이스의 취약점을 점검하고 스캐닝하는 도구의 기반데이터 확보. 플랫폼 기반 데이터를 통해 End 디바이스의 보안 관리 효율성과 편의성을 향상 시킬 수 있을 것으로 기대

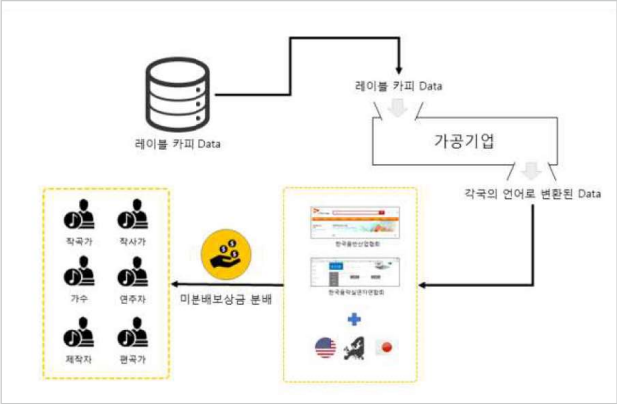
[그림] 개발된 성과물 프로세스 설명 자료

03.
기타

애스토리엔터테인먼트(일반가공)

활용서비스	음원 미분배 보상금 정산 기초데이터 가공
주요내용	국내 음원목록(레이블 카피) 리스트 정리 후 국내외 저작권 단체에 제출 가능한 형태의 데이터로 변환하여 미분배 보상금을 분배 가능한 형태로 가공 ※ 미분배 보상금 발생 원인 : ①보상금 발생 여부를 모르는 아티스트, ②협회 권리자 등록 정보와 일치하지 않음(등록하지 않았거나 매칭 불가), ③보상금 청구를 위한 너무 많은 국가가 있으며, 절차가 복잡함
기대효과	한류, K-POP 등 인기를 얻고 있는 국내 아티스트들의 저작권료 수령이 용이해질 수 있을 것으로 기대 또한, 타 지적재산권 분야(도서, 영상, 그림 등)로 확장가능성이 높음
활용성과	미분배 보상금 신청을 위한 사전 준비가 완료되었으며, 차년도부터 국내외 협회에 미분배 보상금 신청 예정

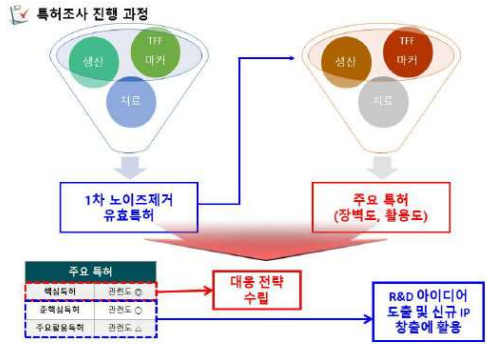
[그림] 데이터 활용 계획



뉴라클사이언스(일반가공)

활용서비스	미국, 유럽의 특허 데이터를 활용하여 자사의 신경계 질환 치료제 기술 개발 전략 수립
주요내용	한국, 일본, 미국, EP, PCT의 신경질환, 항체, 스펙신(Spexin), 펩타이드계 관련 특허 데이터를 분류하여 각 부문별로 구분·분석, 자사 기술개발 방향과 유사한 핵심특허의 분석을 통하여 주요 기술요지 및 발전 동향을 파악
기대효과	2019년 하반기부터 파트너링을 개시하여 2020년 기술이전 또는 공동개발 계약을 체결할 것으로 예상, 2021년에는 글로벌 알츠하이머 치매 임상시험을 개시할 예정으로, 알츠하이머 항체 신약이 발매될 경우, 미국에서만 33조원 매출이 발생할 수 있음 ※ 환자 당 1년 약가 $18,500로 산정, Quontiles IMS Report)
활용성과	스펙신 미국 특허 등록 US10.385.098B2, 신약 후보인 NS100 세포주 개발

[그림] 데이터 활용 성과



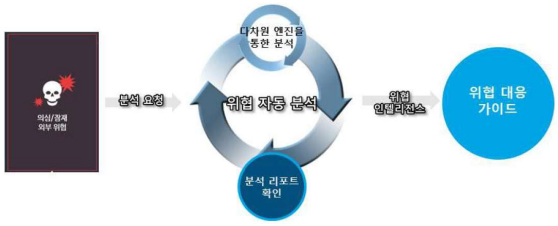
특허 조사 분석 및 연구개발 활용

03.
기타

테이아랩(구매)

활용서비스	악성코드 데이터를 활용한 악성코드 URL 위험진단 서비스
주요내용	악성코드 분석 데이터와 악성 유포지 분석 데이터를 수집 및 분석하여 의심파일 및 의심 URL의 악성 여부를 식별하고, 분류하여 궁극적으로는 종합적인 위험도를 판단하는 CTI제품을 개발 ※ 데이터 활용 단계: ① 1차로 구매한 데이터를 DB화하여 해쉬 검색 및 URL 검색 등의 방법으로 결과를 조회하여 위험을 판단하는 엔진으로 사용 ② 2차로 분석된 결과를 내부 개발자 및 분석가들이 연구하여 자체적으로 분석 할 수 있는 엔진 및 CTI 제품 개발
기대효과	지속적인 연구 개발로 CTI 전문 기업으로 성장하여 보안을 제어하고 지능화된 사이버 공격에 신속하게 대응할 수 있는 가이드 제공
활용성과	다차원 분석을 통한 주요 분석 결과와 안정적인 최신 데이터 수집으로 연구개발의 효율성 증대. 또한, 악성코드를 정확히 분석하고 신속하게 대응할 수 있는 서비스 모델 개발을 통해 선제적 대응 및 실시간 대응이 가능

[그림] 데이터 활용 성과



악성코드를 정확히 분석하고 신속하게 대응할 수 있는 서비스모델 개발